2019

Characterization of genes and genotypes which contribute to lipid accumulation in Saccharomyces cerevisiae and on Zea mays silks

Bri Vidrine
Iowa State University

Characterization of genes and genotypes which contribute to lipid accumulation in *Saccharomyces cerevisiae* and on *Zea mays* silks

by

Bri Vidrine

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics and Genomics

Program of Study Committee:
Marna D. Yandeau-Nelson, Co-major Professor
Basil J. Nikolau, Co-major Professor
Philip W. Becraft
Nick Lauter
Jonathan F. Wendel

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

DEDICATION

This dissertation is dedicated to Dr. Satiander Rana who imparted his magnanimous light on all those around him.

TABLE OF CONTENTS

Page

NOMENCLATURE

ACP	Acyl carrier protein
ANOVA	Analysis of variance
ASI	Anthesis-silking interval
EMS	Ethyl methanesulfonate
ER	endoplasmic reticulum
fao	fatty acid overaccumulation phenotype
GC-FID	Gas chromatography - flame ionization detection
GC-MS	Gas chromatography-mass spectrometry
LC-QTOF-MS	Liquid chromatography - quadrupole time-of-flight mass spectrometry
ORF	Open reading frame
PSE	Post-silk emergence
QTL	Quantitative trait locus
SL	Surface lipid
VLCFA	Very long chain fatty acid

ACKNOWLEDGMENTS

I want to thank my advisors Dr. Marna Yandeau-Nelson and Dr. Basil Nikolau for everything they have to teach me and help me on this journey and I also would like to thank my other committee members, Dr. Nick Lauter, Dr. Philip Becraft and Dr. Jonathan Wendel for their guidance and support.

I want to thank Dr. Satiander Rana for sharing the lipid overaccumulation project presented in chapter two. He oversaw the sequencing and generated all of the backcrossed strains with the help of Pauline Aamodt.

Kori Radke made chapter two and four possible through tireless work and I want to thank her from the bottom of my heart. She helped with data collection for both projects and produced many of the CRISPR strains.

I want to thank Dr. Keting Chen for her role in the development and analysis of the work presented in the fourth chapter.

I want to thank Dr. Jennifer Chmielowski who supported the work presented in the fourth chapter by providing guidance as well as plasmids and *Arabidopisis* strains from her thesis. Dr. Liza Alexander also contributed to this project by helping produce transgenic lines.

I want to thank all the undergraduate students who made this dissertation possible through their hard work. Pauline Aamodt helped produce the backcrossed yeast strains and helped implement the Nile Red assay along with Grant Nickles who ran some of the assays presented in this chapter. Amy Huynh, Zachery Sievers and Emily Naughton helped collect and process data for the fourth chapter, which was a vast undertaking.

Instrumentation and guidance were offered by the W. M. Keck Metabolomics Research Laboratory at Iowa State University managed by Dr. Ann Perera. I want to thank Dr. Lucas

Showman specifically who made significant contributions to this work through his support with method optimization and quantification.

I want to thank my husband, Dr. Ryan Nett, who provided unwavering support through this process both emotionally and intellectually.

I want to also offer my appreciation to the students and teachers who have worked with me. You have all helped make this body of work possible and taught me along the way.

Finally, I would like to thank everyone, including my friends, colleagues, and the department faculty and staff for making my time at Iowa State University a wonderful experience.

ABSTRACT

Lipids are hydrophobic molecules that are essential for life and allow organisms to adapt to changing environments. The chemical properties that make these molecules indispensable for biological activity also make them useful as products that are ubiquitous to human industry and everyday life. Investigating the biological role and biosynthesis of these compounds can reveal new routes for producing efficient and sustainable products and crops. Herein, the study and characterization of lipid biology, genetics, and biochemistry is described in three model systems. Using a forward genetic screen, several genes that impact lipid accumulation in *Saccharomyces cerevisiae* were elucidated. Lipid constituents also coat plant cuticles, creating a barrier between hydrated tissues and the environment. To investigate the protective role of these surface lipids against tissue desiccation, stigmatic maize silks were subjected to temperature-relative humidity combinations and assessed for their rates of water loss across these conditions, while their surface lipid profiles were assessed in tandem to consider association between surface lipids and rates of water loss. Alkane accumulation was found to predict lower rates of water loss, suggesting that alkanes play a protective role against silk desiccation. Because the genes directly involved in alkane biosynthesis within maize have not been fully described, the maize gene, *ZmGl1-like2,* was explored for its potential role in the terminal steps of alkane biosynthesis due to sequence homology to the putative fatty acyl reductase, CER3. Metabolomic characterization in *Arabidopsis cer3* mutants overexpressing *ZmGl1-like2* revealed that *zmGL1-like2* is a functional homolog to CER3 thereby furthering our understanding of alkane biosynthesis in maize, a major cereal crop. Together, this work identifies novel mutations and genes involved in lipid biosynthesis and regulation and considers the relationship between lipid accumulation the underlying biological significance imparted by the lipid metabolome.

CHAPTER 1. INTRODUCTION AND GENERAL BACKGROUND

Overview of lipid biosynthesis

Life depends on lipids. They provide us with energy that can be consumed, stored or burned to power our machines. Within cells, lipids store energy, can serve as signaling molecules, are used to create compartments which separate chemical reactions, and offer protection against the external environment. Lipids are defined as biomolecules that are soluble in organic solvents, but mostly insoluble in water (Moss, Smith, & Tavernier, 1995). Organisms also rely on these hydrophobic molecules to create barriers. These lipid barriers can exist within cells, in the form of organelles, which serve to compartmentalize enzymatic reactions such as in photosynthesis and ATP synthesis. Lipid barriers also exist between cells, in the form of a lipid bilayer, and they can exist on the outside of cells, as the outermost layer of the plant cuticle (Alberts B, Johnson A, 2002; Javelle, Vernoud, Rogowsky, & Ingram, 2011). These assorted functions may correspond to different classes of lipids. Although there is great breadth in chemical diversity among lipids, nature can generate these molecules through subtle modification of simple precursor molecules that are found in all branches of life.

De novo fatty acid biosynthesis

Lipids are produced from fatty acid building blocks that are synthesized through the recursive addition of two carbons to a growing acyl chain. Fatty acid biosynthesis is initiated by the conversion of malonyl-CoA (coenzyme A) to form malonyl-ACP (acyl carrier protein) and fatty acid extension relies on the addition of two carbon moieties derived from malonyl-ACP precursors through a cycle of four iterative reactions: 1. Claisen condensation of acyl-ACP and malonyl-ACP with a concomitant release of CO_2; 2. NADPH-coupled reduction of the beta-ketone to a hydroxyl group; 3. Dehydration of the hydroxyl resulting in a double bond, and 4.

Hydrogenation of the double bond through NADPH coupled reduction (Wakil & Barnes, 1971). This series of reactions is catalyzed in the cytosol by a series of independent enzymes in prokaryotes (e.g. *E. coli*) through fatty acid synthase (FAS) Type II (White, Zheng, Zhang, & Rock, 2005). Plants also utilize FAS Type II within their plastids. Eukaryotic *S. cerevisiae,* however, utilize the FAS Type I complex which consists of two large multi-functional proteins (Lomakin, Xiong, & Steitz, 2007).

Microbial Lipid Accumulation

Efforts to increase fatty acid flux in *S. cerevisiae* to produce oleochemicals

The chemical properties that make lipids essential for many natural processes in living organisms also make these molecules important commodities within industry. Biologically produced oleochemicals (or chemical compounds derived from biological sources) are promising alternatives to petroleum-based products. However, their feasibility as alternatives to fossil fuel-based chemicals hinge upon the cost of production. Unlike oleochemicals derived from plant extracts, microbial hosts have higher potential to be efficiently targeted through targeted engineering. The introduction of enzymes with combinations of broad and/or narrow substrate specificities has the potential to produce novel chemical products that could lead to new markets (Akhtar, Turner, & Jones, 2013; Lennen & Pfleger, 2013). *E. coli* and *S. cerevisiae* systems have been employed to produce oleochemical products including lubricants, soaps and surfactants (Akaike, 1985; P. Liu & Jarboe, 2012). However, low endogenous fatty acid content in these microbial systems is a potential limitation to efficient oleochemical production because fatty acid precursors are necessary for these products; therefore, fatty acid biosynthesis has been a direct target in engineering approaches.

Incremental increases in fatty acid production have been achieved by modifying enzymes directly involved in fatty acid biosynthesis, by introducing genes encoding these enzymes from other organisms, and by removing genes involved in lipid degradation. For example, acetyl-CoA carboxylase (ACC) has been a common target of these efforts because of its rate-limiting role in fatty acid synthesis (de Jong, Shi, Valle-Rodriguez, Siewers, & Nielsen, 2015; Leibundgut, Maier, Jenni, & Ban, 2008; T. Liu, Vora, & Khosla, 2010). Overexpression of ACC was found to increase the production of fatty acids and additional downstream products in conjunction with other overexpression targets such as a fatty acid reductase and deletion of genes in competing pathways such as ADH6 (d'Espaux et al., 2017; de Jong et al., 2015). Disrupting lipid degradation pathways has also been somewhat successful. The triacylglycerol lipases, TGL3, TGL4 and TGL5 have all been targeted to impair the breakdown of triacylglycerols, thus increasing flux in fatty acid biosynthesis (Karin Athenstaedt & Daum, 2005; Ferreira et al., 2018; Runguphan & Keasling, 2014). Like triacylglycerol lipases, beta oxidation of fatty acids has also been a target for removal because of its role in fatty acid degradation. Cumulatively, these engineering approaches have had some success at increasing fatty acid content. However, additional improvements are still required for these systems to be economically viable worldwide and engineered manipulations can interfere with essential lipid metabolism. An improved understanding of the basic mechanisms that underlie the lipid metabolic network may be necessary to achieve this goal.

An alternative to targeted engineering approaches for improving fatty acid production within *S. cerevisiae* is through a forward genetics approach. Forward genetics enables the identification of novel variants as opposed to directed engineering, thus, expanding our genetic understanding of the underlying metabolic processes. Although much is known about the core

machinery of fatty acid biosynthesis, it is often difficult to predict outcomes of genetic manipulations in the broader context of the cell. Therefore, an untargeted genetic screen offers a format to explore this complexity and discover previously uncovered genetic targets to complement ongoing engineering efforts. Ethyl methanesulfonate (EMS) is a point-mutagen which can be utilized to induce untargeted mutations primarily through guanine alkylation, which can result in single nucleotide polymorphisms (particularly transition of cystine to thymine). One EMS-mutagenesis screen identified a novel triacylglycerol lipase, AR1, through the application of a lipase inhibitor assay (Ploier et al., 2013). An EMS-derived mutant from the oleaginous yeast, *Y. lipolytica*, caused 55% higher fatty acid content (L. Liu, Pan, Spofford, Zhou, & Alper, 2015). Sequencing revealed that this mutation resulted in a nonsense mutation within the succinate dehydrogenase, *UGA2*. Therefore, genetic screens can be an effective route for identifying novel targets for enhancing fatty acid production in yeast.

Plant Surface Lipids

The plant cuticle

The plant cuticle is a protective layer that covers aerial surfaces of epidermal cells consisting of the cutin matrix, which is a polymer of ester-linked omega-hydroxy acids and their derivatives, infused and coated with hydrophobic waxes. Epicuticular and intracuticular waxes are epidermal surface lipids that can be extracted with chloroform and and are primarily composed of aliphatic (i.e. non-aromatic) lipids derived from VLCFAs, including hydrocarbons, fatty acids, ketones, alcohols, aldehydes and wax esters (Samuels, Kunst, & Jetter, 2008).

Biological diversity of plant epicuticular surface lipids

The hydrophobic character of surface lipid constituents is utilized by plants to limit transpiration and repel water from surfaces and reduce particulate deposition (Kerstiens, 1996).

In addition to these functions, surface lipids can protect against UV radiation (Long, Patel, Cory, & Stapleton, 2003; Shepherd & Wynne Griffiths, 2006), insect herbivory (Dickinson, Yang, Schuske, & Akam, 1993) and both fungal and bacterial pathogen invasions (Javelle et al., 2011; Kolattukudy, 1985). Plants have surprisingly plastic metabolomes in both the quantity and composition of their lipid constituents, which may underlie diverse protective functionality. Surface lipid compositions are diverse both across plant species and within species; often lipid constituents are specialized between tissue types (M. A. Jenks, Rich, & Ashworth, 1994) suggesting that certain compositions are more beneficial for the function of one tissue over another.

Hydrocarbon biosynthesis in plants

Fatty acids destined for extension past an 18-carbon chain length, must exchange an ACP co-factor for a CoA co-factor, which is catalyzed first by an acyl-ACP thioesterase to release a free fatty acid, followed by a long-chain acyl-CoA synthetase to synthesize the acyl-CoA (Fehling and Mukherjee, 1991). Subsequent elongation in the endoplasmic reticulum resembles de novo FAS within the plastids because of its iterative cycle of condensation, reduction, dehydration and reduction. However, the enzymes are distinct, accepting CoA-bound substrates rather than ACP bound substrates such that they produce very long-chain fatty acid (VLCFA) products (>22 carbon atoms).

The subsequent conversion from VLCFAs to hydrocarbons has not been fully elucidated. Four distinct models have been predicted: 1) Head-to-head condensation of two VLCFAs catalyzed by an aldehyde deformylating oxygenase has been characterized in *Shewanella oneidensis* (Sukovich et al., 2010), but no gene homologs have been identified in eukaryotes; 2) Reduction of the VLCFA to an alcohol followed by dehydration to form even-carbon chain

hydrocarbon products from even-carbon VLCFA precursors, however, there is conflicting evidence for *Vibrio furnissii* (Brown & Shanks, 2012); 3) Decarboxylation of the VLCFA directly to a hydrocarbon was considered the favored plant mechanism in the 1970's, however, a partially purified enzyme extract from *Pisum sativum* showed the release of carbon monoxide instead of carbon dioxide in conjunction with the production of alkanes; and 4) Decarbonylation of an aldehyde through reduction of a precursor VLCFA, which is supported by the production of carbon monoxide in alkane production (Schneider-Belhaddad & Kolattukudy, 2000).

The surface lipids of ten *eceriferum* Arabidopsis mutants were assessed and eight out of the ten were found to have deficiencies in their hydrocarbon composition (Hannoufa, McNevin, & Lemieux, 1993). One of these mutants, *cer1*, showed particularly high deficiencies in alkane accumulation and was postulated to be a decarbonylase. Another one of these eight mutants, *cer3*, was later predicted to convert fatty acyl-CoA to an aldehyde (X. Chen, Goodwin, Boroff, Liu, & Jenks, 2003; Kurata et al., 2003; Rowland, Lee, Franke, Schreiber, & Kunst, 2007). In 2012, a screen by Bernard et al. found that CER1 interacts with CER3 and a cytochrome b_5, CYTB5s. Heterologous co-expression of CER1 and CER3 within yeast was found to reconstitute plant alkane biosynthesis and co-expression of CYTB5s was found to enhance this function (Bernard et al., 2012). Alkane products were detected only when both *CER1* and *CER3* genes were present, giving support to the decarbonylation model for hydrocarbon biosynthesis in *Arabidopsis*. In maize, 30 *glossy* mutants were identified that affect cuticular lipid accumulation and composition (Schnable et al., 1994). Of these mutants, *GLOSSY1* shares 35% sequence identity at the amino acid level with the putative fatty aldehyde decarbonylase CER1, and 62% sequence identify with the putative fatty acyl reductase, CER3. The corresponding loss-of-function cuticular phenotype for *glossy1* mutants, however, had reduced aldehydes and alcohols,

as compared to the *cer1* and *cer3* mutant chemotypes which are deficient in alkanes and aldehydes in the *cer3* mutant (Hannoufa et al., 1993; Sturaro et al., 2005). Therefore, it is unclear whether the hydrocarbon biosynthetic pathway of maize is homologous to *Arabidopsis*.

The role of surface lipids in water maintenance within plant tissues

As a means of protecting against water loss, plants can modulate the composition of extracellular surface lipids (or cuticular waxes) in response to their environment. For example, surface lipid accumulation and composition have been specifically linked to water availability (D. K. Kosma et al., 2009; Shepherd & Wynne Griffiths, 2006). For example, surface alkane accumulation increases in *A. thaliana* in response to drought stress (D. K. Kosma et al., 2009; Shepherd & Wynne Griffiths, 2006). This phenotypic plasticity, or phenotypic response to environmental changes, indicates a unique interplay between the environment and specific surface lipid constituents that may be used to adjust the protective capacity of the cuticle. Plants including maize, rose, tree tobacco, soybean and sesame accumulate higher leaf wax per area after periodic water stress (Cameron, Teece, & Smart, 2006; K. S. Kim, Park, & Jenks, 2007; D. K. Kosma et al., 2009; Premachandra, Saneoka, Kanaya, & Ogata, 1991). These adjustments in response to water deprivation support the inference that cuticle permeability is influenced by environmental cues (Shepherd & Wynne Griffiths, 2006). Some mutations in cuticle-related genes cause lethality under normal conditions but can be recovered when grown under high humidity (Tanaka et al., 2001; Yang et al., 2008), suggesting that compromised cuticle integrity causes susceptibility to desiccation stress. In *Arabidopsis* knockouts of *CER3,* a putative very-long-chain aldehyde decarbonylase involved in the production of alkanes, the mutant plant is sterile unless grown under high humidity conditions, revealing the protective role of surface lipids against pollen desiccation (Tanaka et al., 2001; Yang et al., 2008).

Maize silks as a model to test the hydrocarbon biosynthetic network

Silks are dramatically long stigmatic floral organs in maize that receive pollen and provide conduits for pollen tube elongation and subsequent fertilization of the ovule. Early silk development is characterized by equal rates of cell division and elongation across the silk length. As development progresses, cell division transitions into cell elongation near the point of silk emergence from the husk leaves and helps drive silk emergence (Fuad-Hassan, Tardieu, & Turc, 2008). Reproductively-mature maize silks exist in two microenvironments; the proximal portion is encased by the protective husk-leaves and the distal portion is emerged into the external environment, and exposed to a variety of abiotic stressors, including drought, high temperature and low humidity.

Surface lipid compositions on maize surfaces are distinct between tissues; silk surfaces for example are rich in hydrocarbons (>80% of total silk surface lipids have been measured in the B73 inbred line) but are distinct from alcohol-rich maize leaves (~60% alcohols) that are naturally low (<20%) in hydrocarbons (Perera et al., 2010). A study of 32 inbred lines demonstrated ~10-fold variation in total surface lipids, and this variation also extends to the composition of these surface lipids, including differences between relative abundances of alkanes, alkenes and fatty acids (Dennison et al., 2019; Loneman et al., 2017). Due to the high concentration of hydrocarbons on silk surfaces and the compositional variation, silks are a useful tissue for studying the unresolved maize hydrocarbon biosynthetic network. Another unique aspect of maize silks is their ability to provide water for pollen hydration through cuticle discontinuities (Heslop-Harrison, Heslop-Harrison, & Reger, 1985). This trade-off between water retention and pollen hydration could explain the unique surface lipid composition found on silk surfaces.

Dissertation organization

The following chapters interrogate three facets of lipid metabolism within *Saccharomyces cerevisiae*, *Zea mays*, and *Arabidopsis thaliana* to elucidate genetic and environmental factors that impact lipid biosynthesis. The goals of this work were to 1) identify novel sequence variants that increase accumulation of triacylglycerols in yeast, with the applied goal to engineer efficient microbial factories for the production of precursors for biorenewable chemicals; 2) assess the protective role of lipids on plant surfaces, specifically corn silks against desiccation in hot and dry conditions; and 3) characterize the function of the corn gene, *Glossy1-like2* and its role in the composition of these protective surface lipids in plants. Collectively, knowledge gained from this research will ultimately lead to downstream applications in agriculture and in the production of bio-renewables.

In Chapter 2, a forward genetic screen in the yeast, *Saccharomyces cerevisiae* was employed to identify mutations that contributed to overaccumulation of cellular lipids. Analysis of the genome sequences of seven *fatty acid overaccumulating* strains yielded a candidate list of potentially causal variants that were introduced by genome editing into wildtype yeast. Metabolite profiling revealed that two unique mutations in the genes *TGL3* and *PIB2* caused increases in total fatty acid content. These novel mutations have potential applications in efforts to increase efficiency of oleochemical production.

In Chapter 3, the role of the cuticle, particularly surface lipids, was examined relative to protection against desiccation. The relationship between natural variation in the maize silk surface lipid metabolome and rates of water loss from silk tissue exposed to different combinations of temperature and relative humidity stress were predicted through statistical modelling. Briefly, silk water content was found to predict higher rates of water loss suggesting

that silks are more susceptible to water loss as they become more hydrated. Modelling predicted alkane abundance to be protective against rapid water loss.

In Chapter 4, the function of *ZmGl1-like2* was interrogated via heterologous expression in wildtype and *cer1* and *cer3* mutants of Arabidopsis. Two separate maize alleles were found to complement the *cer3* mutant by increasing aldehyde and alkane accumulation, establishing that *ZmGl1-like2* is the functional homolog of *CER3*, which is hypothesized to function as a fatty acyl reductase in Arabidopsis.

In Chapter 5, major results from this body of research are discussed and put into a broader context. Applications of this research are also discussed, including uses in agriculture and in efforts to engineer oleochemicals.

CHAPTER 2. A FORWARD GENETICS SCREEN IDENTIFIES NOVEL POINT MUTATIONS IN *PIB2* AND *TGL3* THAT CONFER FATTY ACID OVERACCUMULATION IN *S. CEREVISIAE*

Bri Vidrine[1], Satiander Rana[2], Erin Boggess[3], Pauline Ammodt[2], Korinna Radke[2], Basil Nikolau[2], Julie Dickerson[3], and Marna Yandeau-Nelson[2]

[1] Roy J Carver Department of Biochemistry, Biophysics & Molecular Biology; [2] Genetics, Development and Cell Biology, Iowa State University, Ames IA 50011; [3] Electrical and Computer Engineering Departments

Key words: Fatty acid, lipid, TAG, yeast, Nile Red

Abstract

Fatty acid metabolism is critically important for downstream cellular functions, including energy storage, membrane composition and signaling. Fatty acids can also be produced as oleochemicals that serve as precursors for industrial products (e.g. soaps, lubricants, and biofuels) and are promising alternatives for petroleum-based chemicals. A forward genetics screen was employed in the *Saccharomyces cerevisiae* haploid strain, BY4741 to identify mutants with increased neutral lipid content (i.e. triacylglycerols). Fatty acid over-accumulating (*fao*) mutants were identified from an ethyl methanesulfonate-mutagenized population subjected to a two-step screen: buoyant density-based selection followed by identification of high-lipid strains via staining with a lipophilic dye. Seven unique *fao* strains were backcrossed to remove non-causal mutations and their genomes were subsequently sequenced. Of 98 variations relative to the parent strains, 40 corresponded to missense or nonsense mutations within coding sequences. Candidate mutations in seven genes were introduced individually via genome editing into the BY4741 and BY4742 haploid strains and surveyed for their impact on lipid accumulation. Truncation of Pib2p, phosphatidylinositol 3-phosphate-binding protein2, increased total triacylglycerol accumulation

by 2.4-fold within the BY4741 background. Similarly, a G289R substitution within the triacylglycerol lipase, Tgl3p increased triacylglycerol content by ~50% in BY4741 and ~70% in the BY4742 background. This work identifies novel genetic interventions that can be integrated into efforts to engineer *S. cerevisiae* as a microbial oleochemical factory.

Introduction

Oleochemicals are emerging alternatives for chemicals that have been historically derived from petroleum-based feedstocks. Fatty acid metabolic engineering in both bacterial (*E. coli*) and yeast (*S. cerevisiae*) systems has led to the production of oleochemicals commonly sourced from plant oils, such as soaps, lubricants and surfactants (Akaike, 1985; Jarboe et al., 2010). Importantly, these metabolic engineering approaches can produce not only direct or functional replacements for fossil-based chemicals, but also structurally or functionally unique products, providing entry points into potentially new chemical markets (Akhtar et al., 2013; Lennen & Pfleger, 2013; Madhavan et al., 2017).

Because endogenous fatty acid production in *S. cerevisiae* may not be sufficient at an industrial scale, metabolic engineering and synthetic biology approaches have been employed to modulate the overall yield of triacylglycerols (TAG), fatty acids or derivative products by directly targeting key enzymes and substrates in fatty acid biosynthesis (Runguphan & Keasling, 2014; Y. J. Zhou et al., 2016). Numerous interventions have increased flux through fatty acid synthesis by targeting genes within the native pathways or by introducing exogenous genes into the yeast system. Acetyl-CoA carboxylase (*ACC1*) is a common engineering target because it converts acetyl-CoA into the extender molecule, malonyl-CoA in the rate limiting step of fatty acid synthesis (Leibundgut et al., 2008; T. Liu et al., 2010). Overexpression of *ACC1* has increased the production of fatty alcohols, fatty acid ethyl esters, and also fatty acids when co-overexpressed with two fatty acid synthases, *FAS1* and *FAS2* (Y. Chen et al., 2016; d'Espaux et

al., 2017; de Jong et al., 2015; Y. J. Zhou et al., 2016). Furthermore, short chain fatty acid production has been engineered via the targeting and engineering of acyl-ACP thioesterases, and also by disabling the fatty acid beta-oxidation pathway (Leber & Da Silva, 2014). These rational approaches have focused specifically on engineering fatty acid metabolism in *S. cerevisiae* and are limited to the existing knowledge of these pathways.

Forward genetic approaches provide an opportunity to identify genes in perhaps distantly related pathways that might also play a role in the accumulation of fatty acids. Such approaches can be harnessed in conjunction with genome sequencing to expand our fundamental understanding of fatty acid and TAG accumulation in yeast, to ultimately uncover both independent and combinatorial genetic targets for enhancing lipid biosynthesis or accumulation. For example, in *S. cerevisiae* genes associated with lipid content were identified using a transposon-insertion library (Kamisaka et al., 2006), including *SNF2*, *IRA2*, *PRE9*, *PHO90*, and *SPT21*. Similarly, screening of ethyl methanesulfonate (EMS)-induced point mutants in the oleaginous yeast, *Yarrowia lipolytica* and subsequent scale-up of the engineered strain identified a mutant with a 55% increase in fatty acid content, due to a loss-of-function mutation within *UGA2. UGA2* encodes succinate semi-aldehyde dehydrogenase and was not initially known to play a role in lipogenesis (L. Liu et al., 2015).

In this study, we utilized a forward genetics approach to identify novel EMS-induced genetic variants that cause fatty acid overaccumulation in *S. cerevisiae*. Seven resulting fatty acid over-accumulation (*fao*) strains were subjected to genomic sequencing and analysis via breseq to identify putative causal mutations, which were then tested in genome-edited strains. Mutations in a triacylglycerol lipase, Tgl3p, and the FYVE domain-containing phosphatidylinositol 3-

phosphate-binding protein2, Pib2p, increased neutral lipid accumulation and thereby provide new avenues for rational engineering of yeast to produce oleochemicals.

Results

A forward genetic approach was implemented in *S. cerevisiae* to identify strains with increased lipid accumulation. The haploid wildtype strain, BY4741 was mutagenized with the point mutagen ethyl methanesulfonate, which primarily alkylates guanine nucleotides and can ultimately result in the change of a G-C base pair to an A-T base pair. EMS treatment was conducted for either 20 min or 40 min, followed by a two-step screening process to identify candidate *fao* mutants. First, lipid over-accumulating cells were enriched via density gradient-based centrifugation and second, highly buoyant cells were subsequently stained with the lipophilic dye, Nile Red (**Fig. S1**). From the total of 141 colonies selected from the highly buoyant fraction, seven *fao* strains were identified via Nile Red (six from the 40 min and one from the 20 min EMS treatment) (**Table S1**). These strains were named fatty acid overaccumulation (*fao*) mutants. To purify these strains through the selective removal of mutations unassociated with the *fao* phenotype, each of the seven haploid mutants were iteratively back-crossed for five generations (BC5) to the wildtype parental strains (BY4741 and BY4742). Haploid progeny were selected after each cross via Nile Red staining of strains patched on YPD agar plates. Subsequent characterizations were performed on lipid over-accumulating BC5 strains.

Characterization of fatty acid overaccumulation strains

Fatty acid composition was evaluated via gas chromatography (GC)-based profiling of the parental strains, the seven *fao* mutants and the fatty acid over-accumulating strain, *tgl3;tgl4*, a triacylglycerol lipase double knockout mutant. The wildtype BY4741 and BY4742 strains were shown in this study to differ in fatty acid accumulation (**Fig. 1**). The BY4742 parent exhibited

~2.4-fold higher stearic acid accumulation and 1.8-fold higher levels of minor unsaturated fatty acids (excluding fatty acids with 16 & 18-carbon atoms) as compared to BY4741 (**Fig. S2B**). Moreover, minor variations in fatty acid composition were also observed between the two parental strains. For example, BY4741 accumulated 76% while BY4742 accumulated 69% unsaturated fatty acids.

Total accumulation of fatty acids varied amongst the *fao* strains. All mutant strains, with the exceptions of *fao5* and *fao7*, accumulated significantly higher levels of total fatty acids than the original parent strain, BY4741 (**Fig. 1A**). Mutants *fao4* and *fao6* accumulated 1.8- to 2.2-fold higher fatty acid levels as compared to BY4741, and 1.3 to 1.6-fold higher levels relative to BY4742 (**Fig. 1A**). The *fao1, fao2* and *fao3* mutants accumulated between 1.6 to 1.8-fold higher total fatty acids than the parental strain, BY4741. The composition of fatty acids varied only minorly between the parental and mutant strains (**Fig. S2B**). In general, unsaturated fatty acids comprised approximately 65-80% of the total fatty acids and did not exceed the parent-defined range, despite variation amongst the *fao* strains (**Fig. S2B**). Two specific unsaturated fatty acids, palmitoleic and oleic acids, varied across the parental and *fao* strains (**Fig. 1B**). For palmitoleic acid (i.e. C16:1), the strains could be clustered into two groups, one for which the constituent comprised approximately 45% of total fatty acids (parental strain BY4741, and mutant strains *fao3* and *fao5*) and another for which palmitoleic acid comprised approximately 35% of total fatty acids (parental strain, BY4742, and mutant strains *fao1*, *fao2*, *fao4*, *fao6*, and *fao7*; **Fig. 1B**). Oleic acid composition was altered in only one mutant *fao* strain, *fao3*, which accumulated 3.5% and 5% lower levels as compared to BY4741 and BY4742, respectively (**Fig. 1B**). Otherwise, fatty acid compositions were quite similar amongst the *fao* and parental strains.

To complement GC-based analyses, we employed a high-throughput method to approximate lipid content via spectrophotometric analysis of Nile Red-treated yeast cells (Kimura, Yamaoka, & Kamisaka, 2004; Shi, Ji, Siewers, & Nielsen, 2016; Sitepu et al., 2012). We first compared GC-based quantification of total fatty acids with microtiter plate-based spectrophotometric estimation of lipid content using Nile Red-treated cultures of BY4741, BY4742 and *tgl3;tgl4*. A high correlation in detected lipid accumulation was observed between the two methods (R^2=0.93; **Fig. S3**), demonstrating that the lipid phenotypes observed via Nile Red-staining can be attributed to changes in total fatty acid content. Similarly, assessment of the *fao* mutants via the spectrophotometric assay revealed patterns of Nile Red-fluorescence across the evaluated strains (**Fig. S4**) that are similar to the fatty acid accumulation patterns determined via GC-profiling (**Fig. 1**). Namely, *fao1, fao2, fao4* and *fao6*, as well as the *tgl3;tgl4* triacylglycerol lipase double mutant, exhibited significantly higher fluorescence and fatty acid content as compared to the wildtype strains. Therefore, in subsequent experiments the spectrophotometric assay was employed to identify lipid over-accumulating strains, for which lipid composition could then be profiled via analytical instrumentation.

Identification of sequence variation between parental strains, BY4741 and BY4742

Paired-end short-read genomic sequencing was performed on two biological replicates for all *fao* strains and both parental strains, BY4741 and BY4742. The breseq computational pipeline (Deatherage & Barrick, 2014) identified single nucleotide polymorphisms (SNPs) and insertions/deletions (indel variants) relative to the reference genome derived from the S288c strain (Mortimer & Johnston, 1986), but also in comparison to one another (**Table S2**). Some of these sequence variants between BY4741 and BY4742 may underlie the lipid accumulation differences observed between these strains (**Fig. 1**). Of the 53 sequence variations between the parental strains, 22 were unique to the BY4742 strain, whereas 18 were unique to the BY4741

strain relative to the reference genome, S288c. Within BY4742, seven mutations within coding regions were identified in both replicates and lay within *LYS2* (lysine auxotrophic marker), *TKL2* (transketolase2), *YSC83* (encodes a protein of unknown function), *SLA2* (links actin to clathrin), *STE20* (serine/threonine protein kinase), *GPB2* (guanine nucleotide-binding protein subunit beta2), and *ORF Q0182*. The *GPB2* variant codes for a synonymous mutation, however *TKL2*, *YSC83, SLA2* and *STE20* all code for missense mutations, which could potentially influence lipid metabolism. Three of the intergenic alleles within the BY4742 strain and a fourth allele corresponding to BY4741 within the coding sequence of *PSP2* (polymerase suppressor2), were not passed down to any of the backcrossed *fao* strains, despite five backcrosses, suggesting that these alleles may confer a selective disadvantage. Six more coding variants, were identified in both replicates of BY4741. A variant within *MKK1* (mitogen-activated protein Kinase-Kinase1) and a variant within *PPX1* (exopolyphosphatase) coded for nonsynonymous amino acid substitutions. Finally, the auxotrophic markers for methionine (*MET17*) and lysine (*LYS2*) within BY4741 and BY4742, respectively, could impact lipid metabolism as an effect of altered primary metabolism.

Identification of candidate causal mutations via genome sequencing

To identify putative causal mutations in each of the *fao* strains, the genomic sequences were analyzed via the breseq computational pipeline (Deatherage & Barrick, 2014). A total of 151 SNP and indel variants were identified amongst the sequenced strains, of which 98 were unique to one or more of the mutant strains compared to the parents (**Table 1**) and 58 of these were likely due to EMS-mutagenesis based on C to T and G to A changes. A total of 40 genic variants result in either a non-synonymous amino acid substitution or a premature stop codon, 31 of which reflected the characteristic EMS-induced lesion (i.e. G-A or C-T transition).

A subset of unique variants harboring non-conservative amino acid changes (missense and nonsense mutations) were identified for further analysis based on functional annotation of the associated gene and the number of occurrences of a specific mutation within the sequenced strains (**Table 2**). Both the *fao1* and *fao4* mutants harbor the same mutation in the triacylglycerol lipase gene, *TGL3*. The GPI-anchored aspartyl protease gene, *MKC7* harbored the only other unique amino acid substitution identified within *fao1* and was thus selected for characterization. From the unique mutations identified within *fao4*, only the mutations in *TGL3* and *IRS4*, a gene involved in regulating phosphatidylinositol levels, induced non-conservative amino acid changes. In the *fao3* strain, two genes harbored missense mutations, *CDC1,* which is essential for growth and proposed to cleave ethanolamine from a mannose within a glycosylphosphatidylinositol anchor (Losev, Papanikou, OW, & BS, 2008), and the alcohol dehydrogenase, *ADH7*. Both *fao5* and *fao6* harbor a premature stop codon in *PIB2*, a gene encoding phosphatidylinositol binding protein. For *fao6,* two additional missense mutations were pursued in *UFD1* and *GPM1*, which encode polyubiquitin-binding protein and phosphoglycerate mutase, respectively.

Characterization of genome-edited yeast strains for candidate loci

The seven prioritized variants identified from genome sequencing of the seven *fao* mutants (**Table 2**) were individually introduced into the BY4741 and BY4742 parental strains via CRISPR-mediated genome editing. Edited strains were confirmed via sequencing and were phenotyped for total lipid content via the high-throughput Nile Red-based spectrophotometer assay (**Fig. 3**, **S5, S6**). A previous report found that the null *tgl3* mutant elevated triacylglycerols relative to the total lipid content and another study reported a 2-fold increase in triacylglycerols per dry cell weight compared to the BY4741 control strain (Karin Athenstaedt & Daum, 2005; Kurat et al., 2006). Therefore, the amino acid substitution identified in Tgl3p was hypothesized

to cause a similar phenotype with an increase in total lipids due to elevated triacylglycerol content. Introduction of the *tgl3-G289R* allele into the two parent backgrounds induced a 1.6-fold increase in lipid accumulation (**Fig. 2**). In partial concurrence with this result, GC-FID-based analysis confirmed that the *tgl3-G289R* allele induced fatty acid accumulation in the BY4742 background only (**Fig. S6**). Because the *tgl3-G289R* allele did not induce lipid accumulation to the level observed in the *fao1* and *fao4* mutants in which the lesion was initially identified, we also evaluated the impact of the *irs4-A491V* variant, which was also identified in *fao4.* The *irs4-A491V* allele alone did not increase lipid accumulation as compared to the corresponding parental strains. Moreover, a strain carrying both the *irs4-A491V* and *tgl3-G289R* variants did not differ in accumulation as compared to the single *tgl3-G289R* variant strain (**Fig. S7**). These data suggest that only *tgl3-G289R*, and not the relatively conservative alanine to valine amino acid substitution within Irs4p, contributes to the lipid over-accumulation phenotype in *fao4*.

Both *fao5* and *fao6* share the premature stop codon, W468* in Pib2p. The *pib2Δ468* allele introduced into both parental strains increased lipid accumulation by 1.6-fold (**Fig. 2**), and fatty acid accumulation 1.5 and 1.1-fold compared to BY4741 and BY4742, respectively (**Fig. 3**). A second truncated *pib2* allele, *pib2Δ188* that harbors an eight-base pair deletion (positions 537-544 bp relative to GenBank Accession NP_011492) similarly exhibited a 1.6- and 1.1-fold increase in total fatty acid content relative to BY4741 and BY4742, respectively (**Fig. 3**). Comparison of fatty acid content between the *pib2Δ468* and *pib2Δ188* alleles demonstrated that both alleles increased fatty acid content to similar levels relative to the parent strains, with the most pronounced increase within the BY4741 background. Notably, there was no statistical difference in fatty acid accumulation between the *pib2Δ468* and *pib2Δ188* alleles (**Fig. 3**).

Collectively, these results demonstrate that *PIB2* is a negative regulator of fatty acid accumulation.

To further understand the lipid overaccumulation phenotypes caused by the *tgl3-G289R* and *pib2Δ468* alleles, both strains were first assayed with Nile Red and then profiled via LC-QTOF-MS to distinguish between TAG and phospholipid classes. Lipid accumulation estimated by the Nile Red-based spectrophotometric assay was higher on average in both *tgl3-G289R* strains relative to the parent strains, however, these differences were not statistically different than the parent strains (**Fig. 4A**). Increases in estimated lipid accumulation in both *pib2Δ468* strains on the other hand were statistically significant as expected. LC-QTOF-MS-based analysis of the composition of lipid classes in the *tgl3-G289R* strains revealed a 1.6-fold increase in TAG content (**Fig. 4**). ANOVA analysis was conducted to consider the effects of allele (wildtype, *tgl3-G289R* or *pib2Δ468*), parent strain (BY4741 or BY4742) and their interaction on TAG accumulation (**Table S4)**. As expected, TAG accumulation was significantly impacted by the *tgl3-G289R* and *pib2Δ468* alleles. In addition, Diacylglycerols (DAG), which are minor constituents that accumulate at approximately 10% of TAG levels in the parental strains, increased by 1.5-fold in the *tgl3-G289R* (BY4742) strain. For the *pib2Δ468* strains, TAGs accumulated to ~1.9-fold higher levels than the corresponding parental strain, whereas DAG accumulation was not altered. Phosphatidyl serines, phosphatidyl ethanolamines and phosphatidyl inositol levels were not impacted by either variant, even though *PIB2* encodes a phosphatidylinositol (3)-phosphate binding protein.

Interestingly a third, statistically significant mutation within *CDC1,* the *cdc1-G149E* allele, caused a 30% decrease in estimated lipid accumulation, as assessed by the Nile Red-based spectrophotometric assay (**Fig. 3**). This result was especially compelling because the original

strain harboring this mutation in *CDC1, fao3* (BC5) shared this reduced lipid accumulation phenotype. A second allele within *fao3, adh7-G166D* was considered for its potential contribution to lipid overaccumulation, however, it also did not cause a statistically significant difference in fluorescence intensity relative to wildtype in either of the parent strains (**Fig. S6**). Of the two candidates tested from *fao3*, only the *cdc1-G149E* allele exhibited evidence for a role in the lipid overaccumulation phenotype, mirroring the phenotype of the *fao3* mutant.

Finally, functional assessment of the sequence variants identified in *MKC7* within the *fao1* strain and *GPM1* within the *fao6* strain did not reveal differences in lipid accumulation as assessed by the Nile Red-based spectrophotometric assay (**Fig. 2**). Comparison of the *mkc7-G124C* and *gpm1-G95S* genome edited strains as compared to the corresponding parental strains did not reveal statistically significant differences (**Table S3**).

Collectively, the data show that *tgl3-G289R* contributes heavily to the lipid overaccumulation phenotype in both the *fao1* and *fao4* strains, *pib2Δ468* increases lipid accumulation in the *fao5* and *fao6* strains, and *cdc1-G149E* may contribute to changes in lipid accumulation in *fao3*. It remains possible that other sequence variations within these strains may also contribute to the *fao* phenotype.

Discussion

This work provides strong evidence that two unique mutations identified within this mutagenesis screen confer lipid overaccumulation in *S. cerevisiae.* One single amino acid substitution within *TGL3* and a second nonsense mutation within *PIB2* induce lipid overaccumulation.

TGL3 encodes a gene that has previously been shown to act as a triacylglycerol lipase. A null mutant of *tgl3*, was found to have increased triacylglycerol content relative to wildtype (Athenstaedt & Daum, 2005). Therefore, the causative missense mutation within *TGL3* was

originally prioritized as a proof-of-concept variant allele because previous studies indicated that knocking out *TGL3* causes both compositional and overall increases in triacylglycerol content (Athenstaedt, Zweytick, Jandrositz, Kohlwein, & Daum, 1999; Karin Athenstaedt & Daum, 2005). Because Tgl3p functions as a triacylglycerol lipase (Athenstaedt & Daum, 2003), it is thought that these increases are the result of impaired lipid degradation. Therefore, the identification of a *TGL3* missense mutation within this screen initially suggested that this mutation would cause loss of Tgl3p lipase function thereby allowing TAG to over-accumulate. TAG over-accumulation was in fact observed relative to wildtype strains after introduction of the *tgl3-G289R* allele, therefore this amino acid is important for Tgl3p function. Because Tgl3p is a bifunctional enzyme that is responsible for lysophosphatidylethanolamine acetylation in addition to TAG hydrolysis, the G289R mutation could impact either or both functions (Rajakumari & Daum, 2010). The substitution lies within the Patatin-like phospholipase domain, which spans residues 204-297 aa (PFAM: PF01734) as opposed to the other major PFAM domain, a Triacylglycerol lipase domain spanning 70-194aa (PFAM: PF11815). Therefore, this positively charged, hydrophilic substitution at residue 289 may directly interfere with the Patatin-like phospholipase domain. In future studies, other amino acid substitutions at the 289 amino acid position could be compared to determine which substitutions can induce the lipid overaccumulation phenotype.

Lipid overaccumulation observed in response to Pib2p truncation also conferred a lipid overaccumulation chemotype. Pib2p contains a C-terminal FYVE domain that can bind to the signaling lipid, phosphatidylinositol (Pfam Accession: PF01363; residues 448-526). Both Pib2p truncations characterized in this study exclude this domain, fully in the case of *pib2Δ188,* or mostly, in the case of *pib2*Δ*468*. Through GFP-tagging of Pib2p, Kim & Cunningham observed

localization of a truncated GFP-*Pib2Δ155p* (including residues 1-155 aa) to endosomal-vacuolar membranes. However, the slightly shorter truncation of Pib2p (including 1-428 aa) localized to the cytosol, thus determining that FYVE domain is necessary for localization (A. Kim & Cunningham, 2015). Pib2p has recently been shown to be required in combination with the EGO complex to activate TORC1 (Varlakhanova, Mihalevic, Bernstein, & Ford, 2017), which stimulates the production of sphingolipids. Because this previous work suggests that *PIB2* plays a regulatory role, truncations of Pib2p can result in overaccumulation of TAGs, and Pib2p has a domain which may bind phosphatidylethanolamine, PIB2 may be important in lipid regulation. In addition, *pib2* knockouts may be useful in engineering efforts to increase fatty acids in yeast. In the future, differently sized N-terminal and C-terminal truncations could be characterized relative to lipid content and mediation of interactions between EGO and TORC1.

In this study, we observed that the *cdc1-G149E* allele of *CDC1* influenced lipid accumulation. The location of the glycine-to-glutamic acid substitution at residue 149 (relative to GenBank Accession: X81813) matched a previously documented arginine substitution at this position that caused sensitivity to Mn^{2+} and lethality at high temperatures (Paidhungat & Garrett, 1998). Our analysis of *cdc1-G149E* suggests that this allele has the capacity to alter lipid accumulation. Further characterization of this mutation and its effect both on cell growth and lipid accumulation under different environmental conditions, including high and low temperatures, may shed light on this effect.

An unexpected outcome of this study was the identification of lipid accumulation differences between the haploid strains, BY4741 and BY4742. We would expect that these strains would differ in only a few positions in the genome, particularly at the mating type locus (a vs. **α**) and at auxotrophic markers (e.g. *LYS2, MET17*). However, based on our genome

sequence analysis and lipid profiling, we demonstrate that BY4741 and BY4742 are distinct from one another both on a genetic (i.e. 53 nucleotide polymorphisms between the strains) and chemotypic basis. For example, a 3-base insertion within the coding region of PSP2, was present in both biological replicates of BY4741 while it was absent from both replicates of BY4742 and all other BC5 *fao* mutants. Because *fao* strains were produced through mutagenesis followed by iterative selection and backcrossing, the lipid overaccumulation phenotypes observed in the backcrossed *fao* strains could have been due to genomic reshuffling and selection yielding transgressive offspring with unique combinations of existing variants between BY4741 and BY4742. Alternatively, EMS-induced novel mutations may have caused higher lipid phenotypes and were in turn selectively maintained during the sequential backcrossing. Some combination of these two is likely the case. Reiteration of EMS-induced mutations in the wildtype parents demonstrates the importance of these specific mutations on lipid accumulation.

This work identified novel point mutations associated with fatty acid overaccumulation in *Saccharomyces*. Pib2p involvement in lipid accumulation was demonstrated for the first time and provides a new engineering target for increasing lipid content. This study also illustrates how traditional screen can be paired with backcrossing to identify mutations in novel genes or novel mutations in known genes. Further investigation of candidate mutations not addressed in this study will likely result in the identification of additional causative mutations. In some cases, two mutations were maintained in the same *fao* strain through the six rounds of crossing (for BC5) indicating that both provide a selective advantage which could be additive or synergistic (e.g. only beneficial if both are present). Because selection for high lipid phenotypes was applied, these mutations may underlie lipid phenotypes and consideration of these mutations in tandem could be used to identify genes which interact. Regulatory regions (e.g. 5’ UTR) may be an

important missing piece to consider moving forward. These results integrate into engineering efforts by providing potential targets for rational design.

Methods

Ethyl methanesulfonate (EMS) Mutagenesis and density-based selection

Overnight cultures of *S. cerevisiae* strain BY4741 (mating type-**a**) were used for mutagenesis experiments. The optical density of cells was measured at A_{600} and approximately 2 x 10^8 cells/ml were used for each mutagenesis reaction. Cells were harvested by centrifugation at 6,000 x *g* for 5 min and washed twice with 5 ml sterile distilled water. Pelleted cells were resuspended in 0.1 M sodium phosphate buffer, pH 7.0 and centrifuged at 6,000 x *g* for 5 min. Cells were dispensed in 1.7 ml of 0.1 M sodium phosphate buffer. To the cells, 50 µl of EMS (Sigma-Aldrich) was added and incubated along with a non-mutagenized control at 30ºC, with continuous shaking at 180 rpm for either 20 min or 40 min. After incubation the mutagenesis reaction was quenched with 8 ml of 5% sodium thiosulfate. Cells were pelleted at 6,000 x *g* for 5 min and pellets were washed once with sterile water and three times with YPD media. Lastly, cells were resuspended in 1 ml of YPD media and incubated at 30ºC for 4 hr with continuous shaking at 200 rpm for recovery.

Mutagenized cells were subjected to density gradient centrifugation in 200 µl of 1:1 Percoll (GE Healthcare Life Sciences) and 5 M sodium chloride solution, which was loaded gently atop a solution of 9 ml Percoll and 800 µl of 5 M sodium chloride. Samples were centrifuged at 20,000 x *g* at 4ºC for 45 min. The cells with the highest buoyant density (i.e. those at the surface of the solution) were transferred (100 µl) into a sterile microfuge tube. A dilution series of the highly buoyant fraction from the mutagenized culture and from the untreated control were grown on YPD media plates at 30ºC for 24 hr.

Backcrossing

For the purification of non-causative mutations from each mutant strain, lipid overaccumulation mutants (mating type-**a**) were first crossed to the wild type strain, BY4742 (mating type-**α**). Diploid strains were sporulated and resulting tetrads were digested with zymolase (Treco & Winston, 2008) and dissected and separated using a Nikon Eclipse 50i dissection microscope (Nikon Instruments Inc., Elgin, IL), as previously described (Campbell et al., 2019). Haploid mating type was subsequently determined by plating spores on a lawn of cells of the *bar1(sst1)* genotype (Chan & Otte, 1982), which is sensitive to the alpha-mating pheromone and results in a zone of growth inhibition surrounding colonies that are mating type-**a**. Resulting mating type-**α** strains were then backcrossed to BY4741. Each mutant was backcrossed for a total of five generations (i.e. BC5).

Nile Red assays

Selection via YPD agar patch staining

Individual colonies were evaluated via the lypophilic dye, Nile Red followed by either UV-imaging of stained colonies, fluorescence microscopy or a plate-based spectrophotometer assay. For colony staining, cells were patched onto YPD media and grown at 30°C for 24 hr. Cells were resuspended in YPD media, diluted to the same concentration based on an Absorbance (A_{600}) measurement and then 1.3 µl were spotted onto YPD plates in duplicate and incubated for another 24 hr at 30°C. Cell patches on the agar plate were immersed in a 125 µg/ml solution of Nile Red (Acros Organics, USA) dissolved in DMSO for 30 sec. Stained patches were imaged under UV light with a ChemiDoc XRS+ (BioRad) imaging analyzer. Images were scored on a five-point scale relative to a BY4741 low-lipid control and a *tgl3;tgl4* lipase double knockout, high-oil control (i.e. for spot X: 1 if X<BY4741, 2 if X=BY4741, 3 if BY4741<X<*tgl3;tgl4*, 4 if X=*tgl3;tgl4* or 5 if X> *tgl3;tgl4*) as illustrated in **Figure S1**. Each strain was patched twice on

two separate plates, and resulting scores were averaged between both plates. Averaged scores were used to select high-oil spores.

High-throughput microtiter plate spectrophotometer assay

Lipid accumulation was approximated via a 96-well Nile Red-based spectrophotometric assay adapted from (Shi et al., 2016) in a Synergy HT plate reader (BioTek Instruments). Briefly, the cells were distributed in parallel in 96-well clear plates (Corning™ CLS3370) for optical density measurements and 96-well black-walled plates (Corning™ 3915) for fluorescence measurements. Cultures were diluted to 1:20, 1:10, 3:20 and 1:5 in a final volume of 200 µl. For fluorescence measurements, Nile Red dissolved in DMSO was supplemented at a final concentration of 1 µg/ml such that final concentration of DMSO was 0.1%. The optical density was measured at 600 nm absorbance. Nile Red fluorescence was measured with excitation at 540 nm and emission at 600 nm with a Synergy HT plate reader at 'normal' speed after shaking for 30 sec. Cell-free media samples were measured at all three wave lengths and values were averaged to estimate the background at each dilution and subtracted from cell measurements. Paired absorbance and fluorescence intensity values were used to calculate intercepts and slopes to construct standard curves to predict fluorescence intensity at an absorbance equal to 1.

Fatty acid quantification

Primary cultures were grown overnight in YPD media at 30°C with shaking at 250 rpm. The primary culture ($Absorbance_{600}$~0.01) was used to start 50 ml secondary cultures. After 72 hr, $Absorbance_{600}$ was recorded and 7.5 ml were collected and centrifuged at 2,000 *x g*. Resulting pellets were rinsed with deionized water, flash frozen and stored at -80°C until pellets could be lyophilized at -40°C in a FreeZone 4.5 Liter Freeze Dry System (Labconco, MO). Prior to extraction of fatty acids, lyophilized pellets were supplemented with 50 mg of glass beads (Sigma; 425−600µm) and 10 µg of nonadecanoic acid per mg of dry pellet as internal standard.

Derivatization of fatty acids to form fatty acid methyl esters followed an adapted method (Kates, 2010). Briefly, 1 ml of 1N methanolic-HCl was added to each sample. Samples were sonicated for 4 min and then vortexed vigorously for 5 min, followed by incubation at 70°C for 16 hr. After cooling, 1 ml of 0.9% (w/v) NaCl was added to quench the reaction and lipids were extracted twice with 1 ml of heptane. Gas chromatography was performed with an HP-5MS cross-linked (5%) diphenyl (95%) dimethyl polysiloxane column (30 m in length; 0.25-mm inner diameter) using helium as the carrier gas, and an Agilent Technologies series 6890 gas chromatograph, equipped with a model 5873 mass detector (Agilent Technologies). Derivatized extracts were diluted to 10 µg of standard per ml heptane and 1 µl aliquots were injected into the GC via split-less injection, and the GC oven temperature program was as follows: start at 100°C, increase at a rate of 15°C/min to 200°C and hold for 1 min, increase at a rate of 15 °C/min to 320°C and hold for 5 min. Retention indices of lipid metabolites were calculated by calibration with straight chain hydrocarbons ranging in length from 8 to 40 carbons (C8-C40 alkanes calibration standard; Sigma Aldrich). Quantification analysis was performed using the AMDIS software package with aid from the NIST Mass Spectral library (http://webbook.nist.gov/chemistry/) for compound identification. Metabolite abundances were quantified with the integrated signal for methyl nonadecanoate and normalized by dry mass to obtain µmol/mg dry weight.

Extraction and analysis of lipid species via LC-QTOF-MS

Backcrossed strains (BC5) were grown for 72 hr in YPD media at 30ºC with shaking at 250 rpm. Lipids were extracted according to (Kimbara et al., 2013), with modifications. Briefly, 4 ml of culture was harvested by centrifugation at 500 *x g*, washed with diH_20 and lyophilized. Two to four 5 mm zirconia beads were added along with 160 µl of extraction solvent (tert-butyl methyl ether and methanol, 3:1 v/v supplemented with 1 µM of 1,2-didecanoyl-sn-glycero-3-

phosphocholine internal standard) per 1 mg of dried cells. Samples were vortexed briefly followed by the addition of 50 µl of diH_2O per 1 mg of dry cells. Samples were vortexed for an additional 10 min at 2,000 rpm and then incubated on ice for another 10 min. Debris was pelleted at 3,000 x g at 4°C for 10 min and the supernatant was transferred to a new tube and evaporated completely in a Speedvac DNA 120. The dried residue was resuspended with 0.2 ml of 200-Proof ethanol, vortexed again for 10 min at room temperature at 2,000 rpm followed by a 10,000 *x g* centrifugation for 10 min at 4°C. The supernatant was transferred to a GC vial and analyzed via liquid chromatography (Acquity UPLC HSS T3; Waters) and measured with a Quadrupole Time-of-Flight Mass Spectrometer (LC-QTOF-MS).

The LC program utilized Solvent A, acetonitrile:water:1 M ammonium acetate: formic acid (158 g: 800 g: 10 ml: 1 ml), and Solvent B, acetonitrile :2-propanol:1 M ammonium acetate: formic acid (79 g: 711 g: 10 ml: 1 ml), and was programmed as follows: 35% B at 0 min, 70% B at 3 min, 85% B at 7 min, 90% B at 10 min, 90% B at 12 min, 35% B at 12.5 min with a flow rate of 0.15 ml min^{-1} at 55°C. Samples were measured in both positive and negative ion modes, however, only the negative mode was analyzed because it had a better representation of lipid species. Metabolite classes were identified using Agilent MassHunter software in conjunction with the NIST17 mass spectral database. Quantification was calculated relative to the 1,2-didecanoyl-sn-glycero-3-phosphocholine internal standard and normalized by the dry cell mass.

Genome sequencing and analysis

For whole genome sequencing, yeast DNA was isolated via the Universal DNA extraction kit (Promega, USA) according to manufacturer instructions, and subjected to 150-cycle, paired end sequencing using the Illumina HiSeq 2500 platform at the DNA Facility at Iowa State University. High-throughput sequence data preprocessing, alignment and analysis was conducted using default settings on the cloud-computing platform, CyVerse (Merchant et

al., 2016). Read quality was measured using FastQC (bioinformatics.babraham.ac.uk). The sequenced genomes of mutant strains were aligned with the available *S. cerevisiae* S288C reference genome (Engel et al., 2014) using Bowtie (Langmead, Trapnell, Pop, & Salzberg, 2009). Because yeast is a relatively small microbial genome (~12 Mb) and EMS mutagenesis enriches for point mutations, the breseq computational pipeline (Deatherage & Barrick, 2014) was used to identify variants relative to the reference genome S288C, including SNPs and indels.

CRISPR/Cas9 mediated genome editing

Single site mutations in each strain were confirmed via Sanger sequencing of PCR amplified DNA fragments. Specific mutations were engineered into wildtype BY4741 and BY4742 strains via CRISPR/Cas9 genome editing, using the pCRCT vector that encodes iCas9 and tracrRNA, which was a gift from Dr. Huimin Zhao (Addgene plasmid #60621; Bao et al., 2015). Each guide RNA (gRNA) was identified using the Chop-Chop web server (http://chopchop.cbu.uib.no/). DNA sequences were designed from gRNA sequence as per Bao et al. (2015) to simultaneously code for the gRNA and provide a template for the desired mutation flanked by homologous DNA regions so this fragment was available for homologous recombination after targeted cutting. BsaI restriction sites were also introduced to facilitate in-frame cloning into the pCRCT plasmid to produce the CRISPR/Cas9 construct *in vivo* (Xing et al., 2014). This entire sequence was synthesized as a gBlock (**Table S5**) by Integrated DNA Technologies (Coralville, IA) and cloned into the pCRCT vector via single-step BsaI digestion and T4 ligation. The reaction mixture was directly transformed into *E. coli* (DH5α), selected with ampicillin and confirmed via Sanger sequencing. Plasmids were subsequently transformed into strains BY4741 and BY4742 using the lithium acetate/single-stranded carrier DNA/PEG method (Gietz & Schiestl, 2007). After selection on synthetic media deplete of uracil, colonies were sequenced with targeted primers outside of the plasmid sequence (Table S5) to identify

successful integration of the mutation into the genome. To remove pCRCT, strains were grown without selection on YPD and single colonies were screened for loss of the plasmid by their inability to grow on SD-URA. Mutations introduced by genome editing were confirmed in strains lacking the plasmid via bi-directional sequencing. Double mutant strains were produced by introducing second mutations into edited strains harboring one introduced sequence variation and lacking the pCRCT plasmid through the same process described above.

Statistical Analysis

All statistical analyses were performed using JMP Pro (version 12.0.1, SAS Institute Inc., Cary NC) with a minimum of three biological replicates. Dunnett's tests were used to compare mutant strains with a given control strain (i.e. either BY4741 or BY4742). Two-way ANOVA were conducted on appropriately partitioned datasets to evaluate the effects of wildtype background (i.e. BY4741 or BY4742), and specific mutations on a given trait.

Acknowledgments

We thank Dr. Lucas Showman of the W.M. Keck Metabolomics Laboratory at Iowa State University for assistance with development of LC-QTOF-MS methods, and related data analysis and lipid identification. This study was supported in part by the National Science Foundation Engineering Research Center for Biorenewable Chemicals (CBiRC), Award No. EEC-0813570. We thank CBiRC for the resources that permitted this collaborative work.

Figures

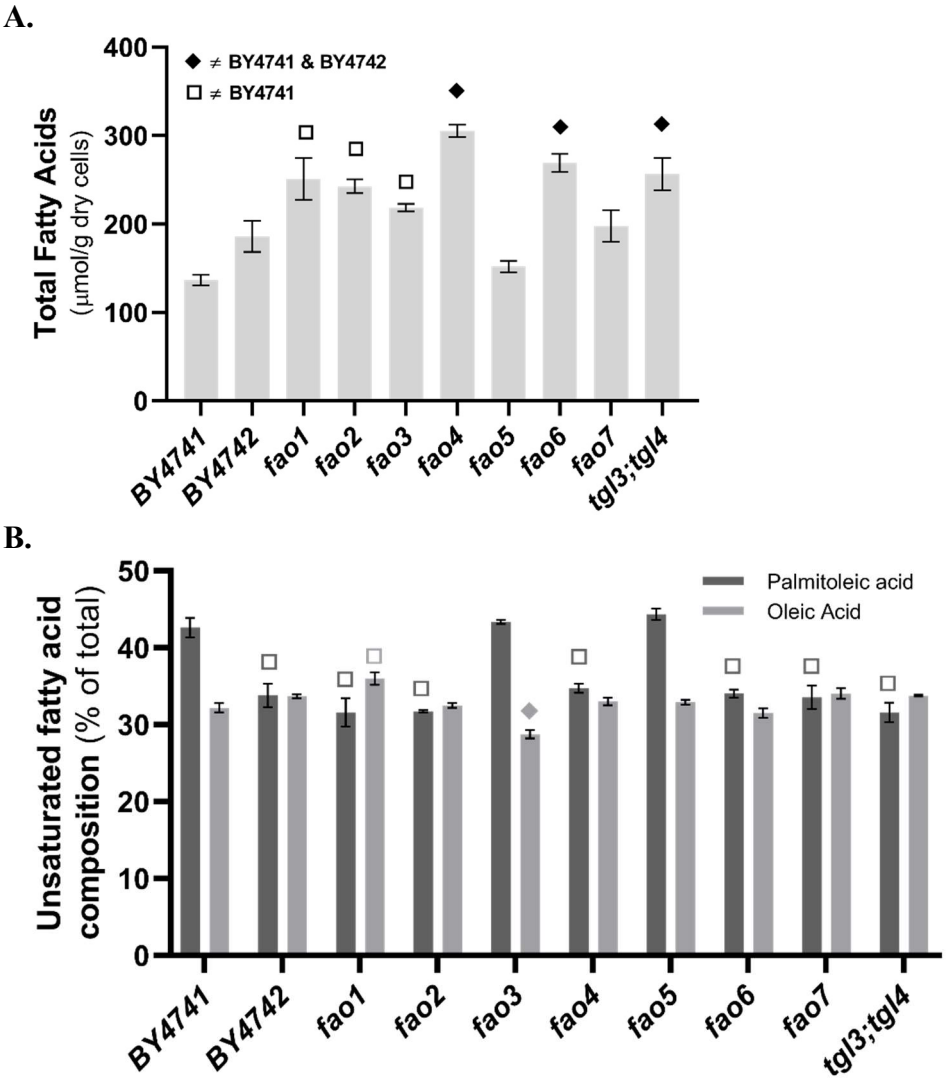


Figure 1 | Total accumulation and relative composition of fatty acids in *fao* mutants. (**A**) Concentrations of fatty acids from *fao* mutants (BC5), wildtype parents and the TAG-accumulating triacylglycerol lipase double mutant, *tgl3;tgl4*. (**B**) Relative abundances (%) of long-chain, unsaturated fatty acid constituents comprising greater than 2% of total fatty acid content. Symbols represent statistical differences (p-value<0.05) relative to BY4741 and relative to both BY4741 & BY4742 according to a Tukey's HSD test. Means ± SE; three biological replicates.

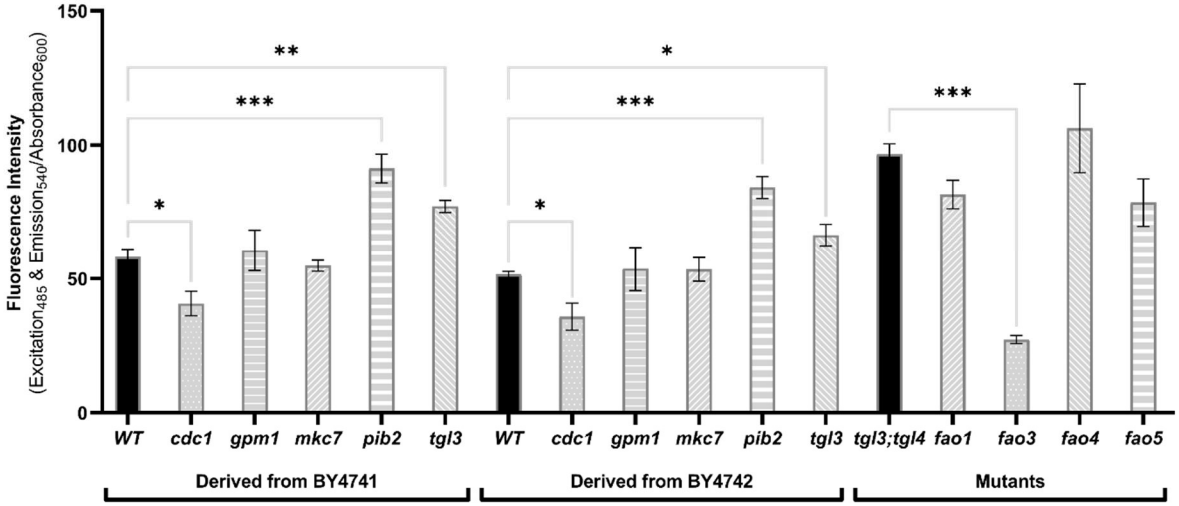


Figure 2 | Lipid accumulation in genome edited lines estimated by Nile Red fluorescence spectrophotometric assay. Genome-edited BY4741 and BY4742 strains harboring candidate point mutations were evaluated for lipid content using the high throughput Nile Red spectrophotometric assay. Fluorescence intensities were normalized by culture absorbance at 600 nm. Estimated lipid content was compared among genome-edited strains and *fao* mutants (BC5) compared to the BY4741, BY4742 and the *tgl3;tgl4* lipid over-accumulating strains via a Dunnett's test. Statistical differences are indicated with asterisks ($p<0.05$*; $p<0.005$**; $p<0.0005$***). Means ± SE for BY4741(n=20) and BY4742 (n=18) parental strains (WT), the genome-edited strains (n=3-7) and the *fao* (BC5) mutants (n=3-7).

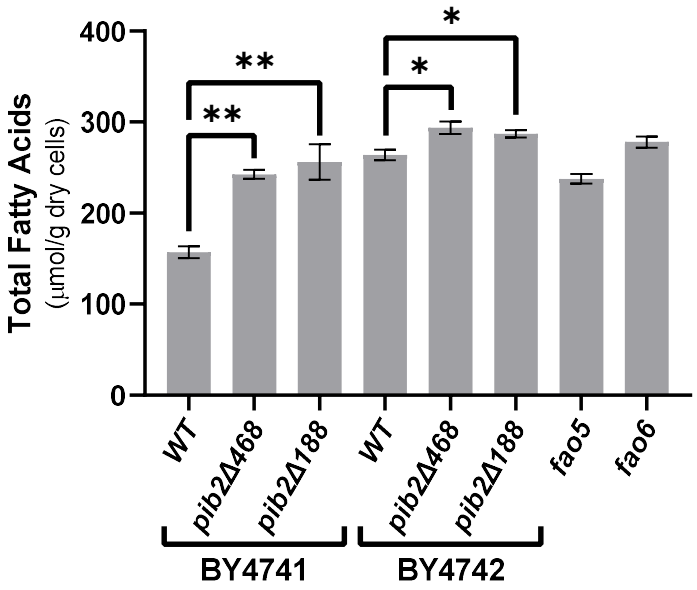


Figure 3 | Total fatty acid accumulation in strains harboring premature stop codon mutations in *PIB2*. Total fatty acid abundances were quantified via GC-FID from strains grown in rich media for 72 hr. Total fatty acid content was compared between the genome-edited strains and the corresponding parental strain via Dunnett's tests. Statistical differences are indicated with asterisks ($p<0.05$*, $p<0.005$**). The original *fao* strains, *fao5* and *fao6* (BC5) strains were grown alongside the corresponding edited strains for reference. Means ± SE; three biological replicates.

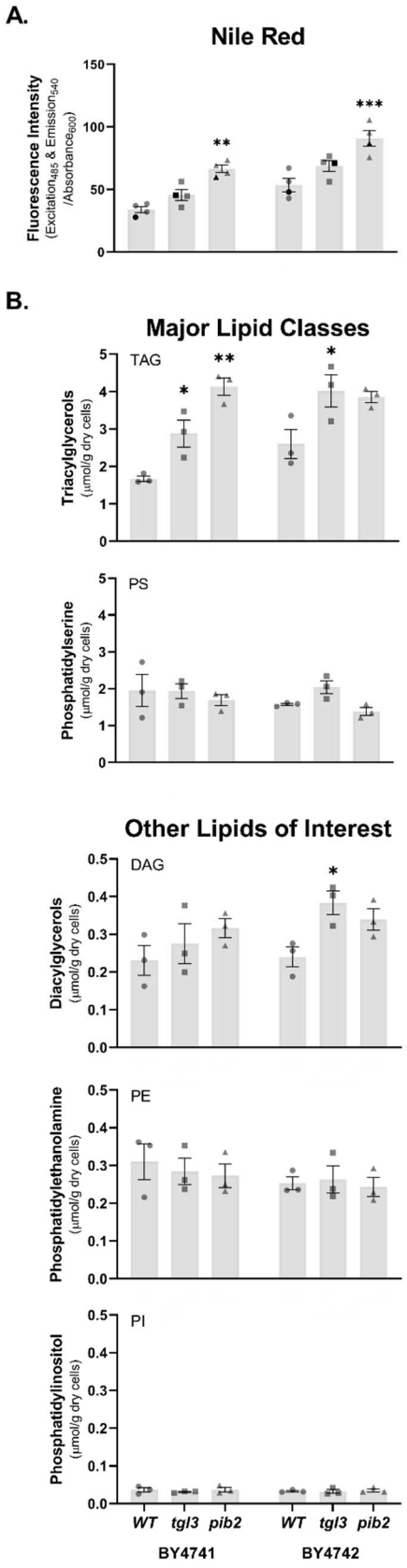


Figure 4 | Comparison of lipid profiles from genome-edited *tgl3* and *pib2* and parent strains. Both parent strains were edited with two different mutations, *tgl3-G289R* (squares) and *pib2Δ468* (triangles) and compared to the respective unedited parent strain (circles). These strains were grown for 72 hr at 30°C in synthetic media. (**A**) Lipid accumulation estimated using the Nile Red-based spectrophotometric assay for four biological replicates. Three of the four replicates from each strain (grey points in panel A) were also measured with LC-QTOF (**B**) to determine abundances of different lipid classes, specifically triacylglycerols (TAG), phosphatidylserines (PS), diacylglycerols (DG), phosphatidylethanolamines (PE), and phosphatidylinositols (PI). Statistical differences between the parent strain and each mutant were determined using a Dunnett's test within each parent background and are indicated with asterisks ($p<0.05$*; $p<0.005$**; & $p<0.0005$***). Means ± SE.

Tables

Table 1. Classification of the 151 variant loci identified via breseq analysis.

Classes of variants			Total Variants	Derived from EMS[a]	BY4741	BY4742	*fao1*	*fao2*	*fao3*	*fao4*	*fao5*	*fao6*	*fao7*
BY4741 *versus* BY4742			53	0	34 (21)[b]	34 (21)	34 (29)	31 (31)	28 (24)	33 (29)	31 (17)	32 (15)	33 (25)
Unique variants within *fao* mutants	Intergenic (e.g. ncRNAs)		37	14	0	0	9 (3)	6 (5)	6 (4)	6 (4)	9 (3)	7 (4)	6 (2)
	Unique mutations within coding regions	Nonsense mutations	4	3	0	0	0	1 (0)	0	0	3 (2)	1 (1)	0
		Missense mutations	36	28	0	0	2 (1)	4 (3)	5 (2)	3 (3)	11 (9)	3 (1)	11 (3)
		Nucleotide deletions	3	0	0	0	1 (1)	1 (0)	1 (1)	1 (1)	0	0	0
		Nucleotide insertion	1	0	0	0	0	0	0	0	0	0	1 (0)
		Nucleotide slippage (e.g. (G)9→(G)10)	1	0	0	0	0	0	0	0	1 (0)	0	0
		Silent mutations	16	13	0	0	1 (1)	4 (4)	3 (0)	0	4 (4)	1 (1)	4 (0)

[a]Mutations which are consistent with EMS induced mutationsbThe cumulative number of variants from both sequenced replicates and in parentheses, the mutations which were present in both replicates.
[b]The cumulative number of variants from both sequenced replicates and in parentheses, the mutations which were present in both replicates.

Table 2. The subset of mutations within candidate genes that were introduced via genome editing into the BY4741 and BY4742 strains.

Strains	Mutation	Annotation of mutation	Genes	Gene Description
fao1	C→A[a]	G124C (GGC→TGC)	*MKC7*	Aspartyl protease
fao1 *fao4*	C→T	G289R (GGG→AGG)	*TGL3*	Triglyceride lipase/lysophosphatidyl-ethanolamine acyltransferase
fao3	G→A	G149E (GGG→GAG)	*CDC1*	Putative mannose-ethanolamine phosphate phosphodiesterase
fao3	G→A	G166D (GGT→GAT)	*ADH7*	NADP-dependent alcohol dehydrogenase
fao4	G→A	A491V (GCA→GTA)	*IRS4*	Regulates phosphatidylinositol 4,5-bisphosphate
fao5 *fao6*	C→T	W468* (TGG→TAG)	*PIB2*	Phosphatidyl inositol binding protein
fao6	C→T	G95S (GGT→AGT)	*GPM1*	Phosphoglycerate mutase

[a]Not predicted by EMS

*Asterisk indicates a stop codon.

Supplement

Supplemental Figures

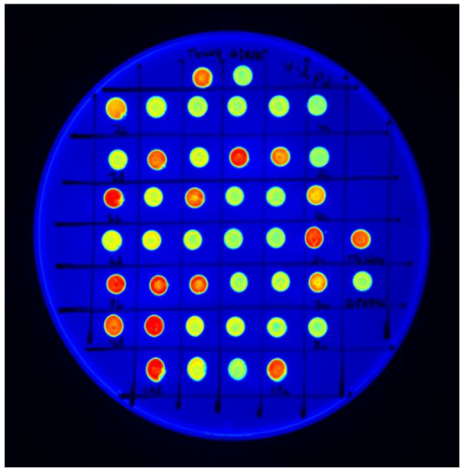


Figure S1 | Selection of backcrossed spores via Nile Red-staining of patches on rich media agar. Individual spores spotted onto rich media agar plates were stained with the lipophilic dye, Nile Red. Stained yeast strains were visualized under UV light and imaged. Individual strains were spotted on two separate agar plates and scored twice. Scores were based on comparisons relative to the BY4741 parental strain and the double triacylglyceride lipase mutant *tgl3;tgl4*, which over-accumulates triacylglycerols. Strains that fluoresced less intensely than BY4741 were scored as '1,' strains that fluoresced equally as BY4741 were scored as '2,' if they were between BY47412 and *tgl3;tgl4* they were scored as '3,' if strains were equal to *tgl3;tgl4* they were scored as '4,' and if they exceeded *tgl3;tgl4* they were scored as '5.' Averaged scores were used to select high-lipid spores for further backcrossing.

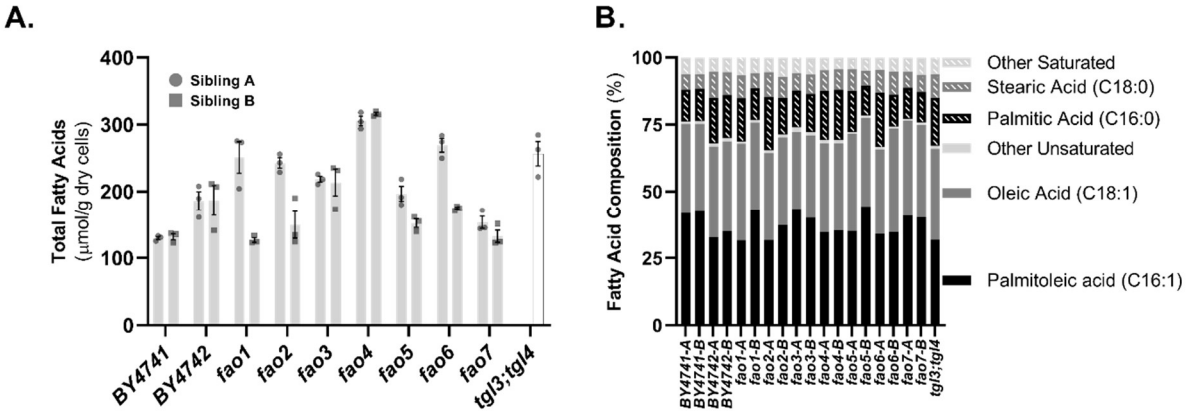


Figure S2 | Fatty acid analysis of all backcrossed *fao* mutants subjected to genomic sequencing. Two sibling BC5 strains, sibling A (circles) and B (squares) were selected for each of the seven *fao* mutants and the two wildtype parents, and subjected to genomic sequencing. **(A)** Total fatty acids were quantified via GC-FID. (**B**) The composition of major saturated (striped bars) and unsaturated (solid bars) fatty acids are represented as percentages. In both panels, high-triacylglycerol *tgl3;tgl4* strain is shown. Means ± SE; three biological replicates.

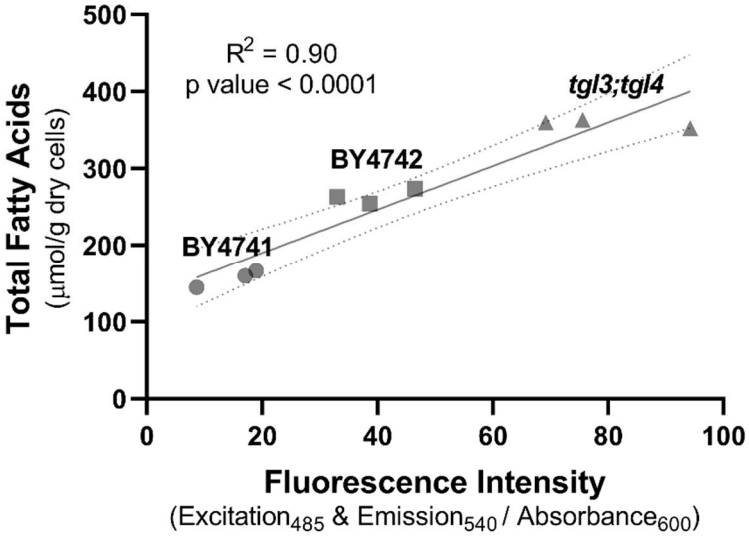


Figure S3 | Predictive power of high-throughput microtiter plate spectrophotometer assay for total fatty acid accumulation. Individual cultures grown in rich media for 72 hr were assayed via the Nile Red 96-well plate spectrophotometric assay and extracted fatty acids were profiled via GC-FID. Fluorescence intensity in the spectrophotometric assay and fatty acid accumulation values are plotted for each strain. Linear regression shows strong association between the two methods.

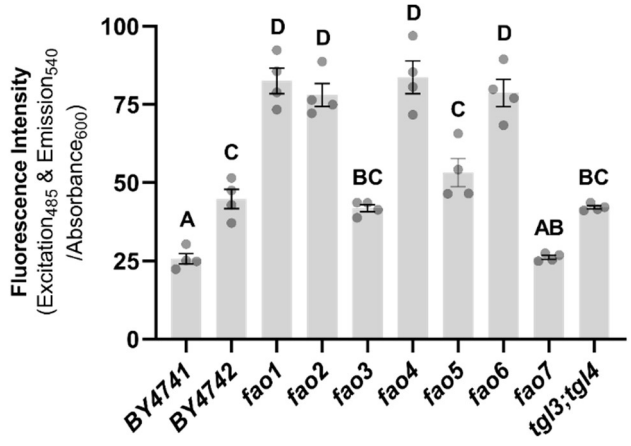


Figure S4 | Lipid accumulation in the *fao* BC5 strains estimated by Nile Red fluorescence spectrophotometric assay. Sibling A from each of the sequenced BC5 mutants was grown in synthetic media within a 96-well plate and subjected to the Nile Red fluorescence spectrophotometric assay. Fluorescence values were normalized relative to absorbance of the culture. Different letters above data-bars denote statistically significant differences at $p<0.05$, according to a Tukey's HSD test. Similar results were obtained in a second replication of the experiment. Means ± SE; four biological replicates.

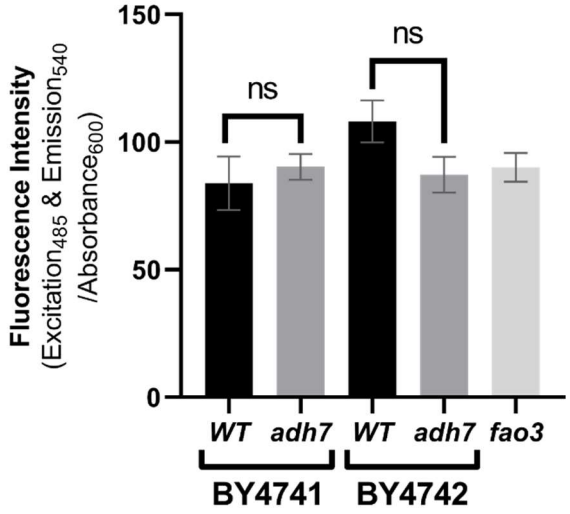


Figure S5 | Lipid accumulation in *adh7* mutants estimated by Nile Red fluorescence spectrophotometric assay. The genome-edited strains (*adh7-G166D*), parental strains and the relevant *fao* strain, *fao3* were grown in rich media and subjected to the Nile Red fluorescence spectrophotometric assay. Nile red fluorescence was compared between the genome-edited strains and the corresponding parental strain via Dunnett's tests. No statistically significant differences were detected ($p>0.05^{ns}$). Means ± SE; seven biological replicates.

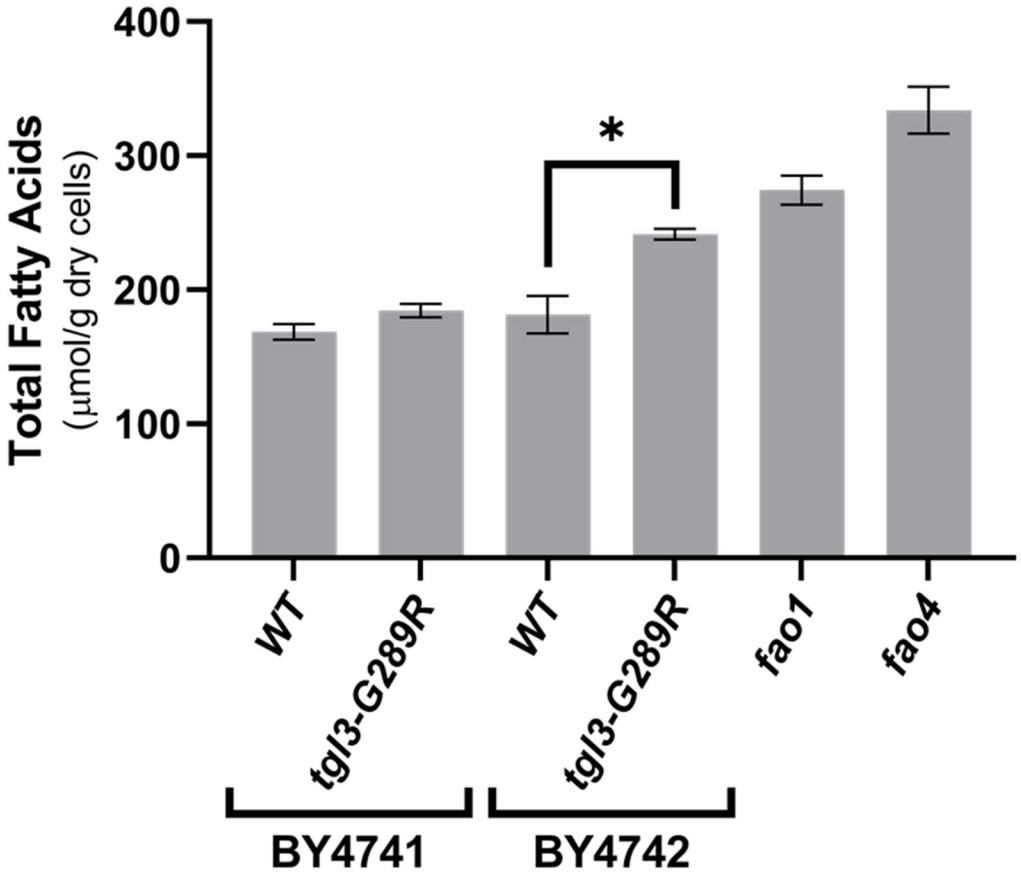


Figure S6 |Fatty acid accumulation in the genome-edited *tgl3-G289R* strain. Total fatty acid abundances quantified via GC-FD from *tgl3-G289R* strains grown in synthetic media. Total fatty acid content was compared between the genome-edited strains and the corresponding parental strain via Dunnett's test. Statistical differences are indicated with asterisks ($p<0.05$*, $p<0.005$**, $p<0.005$***). The *fao1* (BC5) and *fao4* (BC5) strains were grown alongside corresponding genome-edited strains for reference. Means ± SE; three biological replicates.

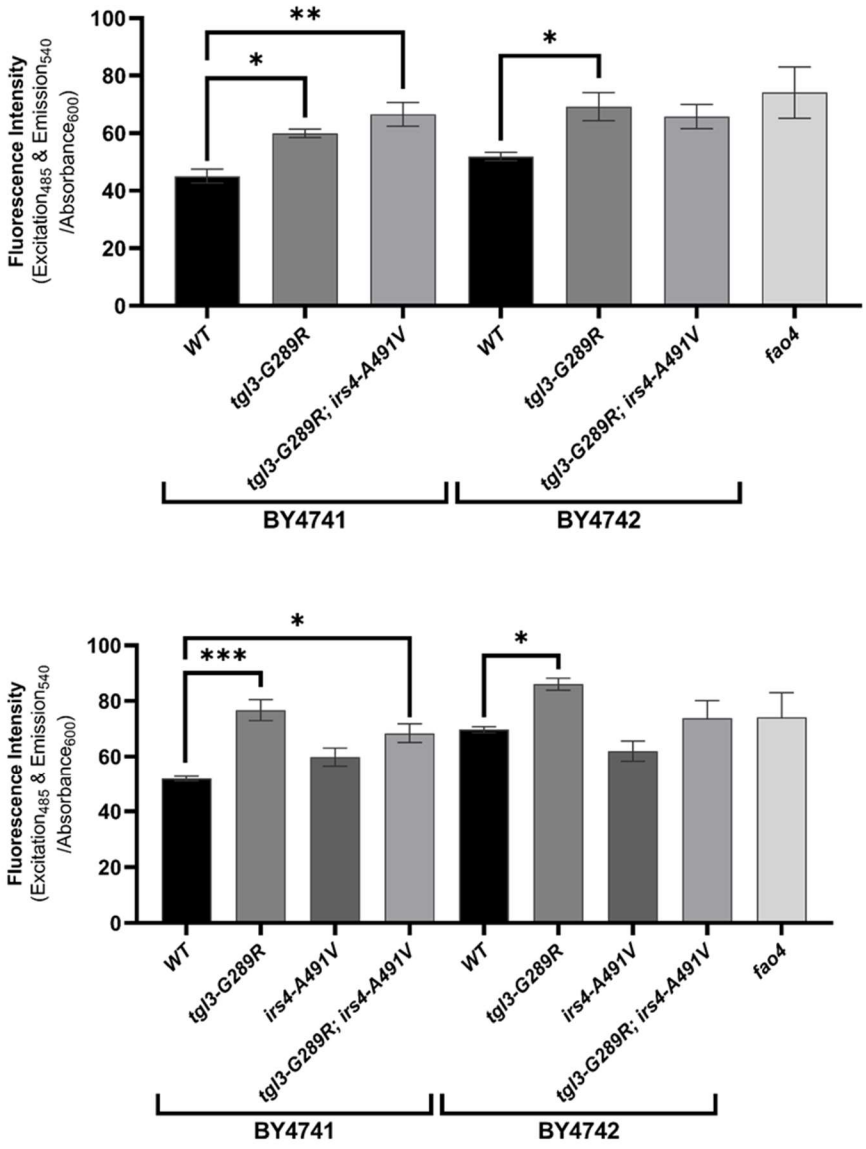


Figure S7 | Lipid accumulation in genome-edited *tgl3-G289R, irs4-A491V*, and *tgl3-G289R; irs4-A491V* mutants estimated by Nile Red fluorescence spectrophotometric assay. The genome-edited strains (*tgl3-G289R, irs4-A491V*, and *tgl3-G289R; irs4-A491V*), parental strains and the relevant *fao* strain, *fao4* were grown in rich media in a 96-well plate for 48 hr and subjected to the Nile Red fluorescence spectrophotometric assay. Nile Red fluorescence was compared between mutants and wildtype strains within each parent background. Strains were measured (**A**) after plasmid removal and (**B**) while still harboring plasmid containing the CRISPR construct. Nile Red fluorescence was compared between the genome-edited strains and the corresponding parental strain via Dunnett's tests. Statistical differences are indicated with asterisks ($p<0.05$*, $p<0.005$**, $p<0.005$***). Means ± SE; four biological replicates.

Table S1. Recovery of one and six lipid over-accumulating *fao* strains after two different durations of EMS treatment.

Exposure Time	Colonies Screened	Recovery	Confirmed
Control (0 min)	39	0%	0
20 min	70	14%	1
40 min	71	85%	6

Table S2. Variants identified between BY4741 and BY4742 strains. Variants relative to the S288c reference genome are indicated with a '1.' Out of these variants, those which were detected in only the BY4741 parent (green) or only the BY4742 parent (yellow) are shaded with green and yellow respectively. Variants identified within at least one replicate of both parents are unshaded.

Variant location		Variants identitified relative to the reference genome, S288c, are indicated by a 1.																				Variant information		
Chrom	Pos	BY4741-1	BY4741-2	BY4742-1	BY4742-2	fao1-1	fao1-2	fao2-1	fao2-2	fao3-1	fao3-2	fao4-1	fao4-2	fao5-1	fao5-2	fao6-1	fao6-2	fao7-1	fao7-2	Variant	Annotation	Gene Name
AJ011856	33789			1	1															(C)10→9	intergenic (+2775/-1584)	ORF Q0092 → / → tE(TTC)Q
AJ011856	33800			1	1		1	1	1			1	1	1		1		1		(G)7→8	intergenic (+2786/-1573)	ORF Q0092 → / → tE(TTC)Q
AJ011856	35803	1	1								1				1					(T)9→10	intergenic (+359/-737)	tE(TTC)Q → / → CYTB
AJ011856	36306			1	1															(A)9→8	intergenic (+862/-234)	tE(TTC)Q → / → CYTB
AJ011856	52707	1	1			1									1	1	1		1	(AATTAT)3→4	intergenic (+1278/-1402)	ORF Q0143 → / → ORF Q0144
AJ011856	63666			1	1		1	1	1			1	1	1				1		(G)10→9	intergenic (+1219/-196)	21S ribosomal RNA (R2) → / → tT(TGT)Q1
AJ011856	65926			1	1		1	1	1			1	1	1		1	1	1		(G)9→8	coding (157/405 nt)	ORF Q0182 →
AJ011856	66837	1	1			1									1				1	C→G	intergenic (+552/-224)	tQ(TTG)Q → / → tK(TTT)Q
AJ011856	67409			1	1	1	1		1			1	1	1		1	1	1		(C)9→10	intergenic (+28/-59)	tR(TCT)Q1 → / → tG(TCC)Q
AJ011856	82367			1	1		1	1	1			1	1	1		1				A→C	intergenic (+2345/-2668)	COX3 → / → tM(CAT)Q2
BK006934	22	1		1	1	1		1	1				1		1			1	1	+C	intergenic (–/+423)	– / ← YHL050C
BK006934	2303	1	1	1		1	1	1		1	1		1	1	1	1	1		1	(C)11→12	intergenic (-406/+368)	YHL050C ← / ← YHL050C
BK006934	95446			1	1		1	1											1	G→A	T831I (ACA→ATA)	STE20 ←
BK006934	138918			1	1		1	1	1			1	1	1	1	1			1	T→A	H75Q (CAT→CAA)	YSC83 →
BK006934	499958	1	1			1				1	1			1	1	1	1	1	1	C→T	E396K (GAA→AAA)	PPX1 ←
BK006934	511255	1	1			1				1	1			1	1	1	1	1	1	C→T	F631F (TTC→TTT)	SCH9 →
BK006935	6755	1	1			1	1			1				1					1	(A)19→20	intergenic (+4048/+480)	YAL067W-A → / ← SEO1
BK006935	27060		1		1	1	1	1	1	1		1		1						A→T	V303V (GTT→GTA)	FLO9 ←
BK006935	27063				1	1		1	1	1										G→A	G302G (GGC→GGT)	FLO9 ←
BK006935	27075		1		1	1	1	1	1	1		1		1						G→A	T298T (ACC→ACT)	FLO9 ←
BK006935	27084		1	1	1	1	1	1	1	1		1		1	1	1	1			G→A	T295T (ACC→ACT)	FLO9 ←
BK006935	27129				1		1	1		1				1						C→G	T280T (ACG→ACC)	FLO9 ←
BK006935	41475			1	1		1	1	1			1	1			1				G→T	V739V (GTG→GTT)	GPB2 →
BK006936	5983			1	1	1	1	1	1			1	1		1				1	C→G	intergenic (-974/+1622)	YBL111C ← / ← PAU9
BK006936	469725			1	1		1	1	1							1	1	1		4495 bp→6 bp		LYS2
BK006936	474225			1	1		1	1	1							1	1	1		C→A	intergenic (-299/+167)	LYS2 ← / ← TKL2
BK006936	474242			1	1		1	1	1							1	1	1		Δ1 bp	intergenic (-316/+150)	LYS2 ← / ← TKL2
BK006936	474544			1	1		1	1	1							1	1	1		A→G	F632L (TTC→CTC)	TKL2 ←
BK006937	123460			1	1							1						1		A→G	intergenic (-457/+117)	YCR006C ← / ← SUF2
BK006938	548565				1	1	1	1	1		1		1		1	1				(T)19→20	intergenic (-255/+197)	RPC11 ← / ← BAP3
BK006938	688076	1		1	1	1	1	1		1	1		1	1	1	1	1		1	(T)14→15	intergenic (+236/-151)	APC4 → / → VBA4
BK006938	688076		1															1		(T)14→16	intergenic (+236/-151)	APC4 → / → VBA4
BK006938	769328			1	1		1	1	1				1		1				1	G→A	intergenic (-328/-197)	CPR1 ← / → RPA14
BK006938	1254955	1	1							1				1	1	1	1	1	1	A→G	D631D (GAT→GAC)	UBA2 ←
BK006938	1518961	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		(T)13→14	intergenic (-1286/+703)	IRC4 ← / ← YDR541C
BK006940	106290		1			1		1			1		1	1	1	1	1	1	1	(T)13→14	intergenic (-54/+179)	MDJ1 ← / ← YFL015C
BK006940	210406			1															1	(TG)31→33	intergenic (-338/+213)	CDC14 ← / ← SUP6
BK006941	496742			1		1	1	1		1			1	1	1				1	(A)13→14	intergenic (-240/+391)	ERG26 ← / ← EFM5
BK006941	530056	1		1					1		1	1					1	1	1	+84 bp	coding (793/1656 nt)	MTL1 →
BK006941	531875	1																		(TA)15→16	intergenic (+194/+8)	tD(GUC)G1 → / ← THG1
BK006945	13636			1	1															C→G	intergenic (-191/+1012)	PAU18 ← / ← AYT1
BK006945	732283	1	1			1								1	1		1	1		2403 bp→5 bp		[YLR302C]–MET17
BK006945	734703	1	1			1								1	1		1	1		T→A	intergenic (+827/-99)	MET17 → / → tI(AAU)L1
BK006946	184075			1	1		1					1	1							G→T	intergenic (-107/-386)	tG(GCC)M ← / → YML045W
BK006946	238324	1	1																	(ACA)7→8	coding (1371/1778 nt)	PSP2 →
BK006946	273138	1	1							1	1			1		1		1	1	T→G	G7G (GGT→GGG)	AIM34 →
BK006947	12986		1	1	1		1				1		1		1					(T)11→12	intergenic (-110/-281)	SNO2 ← / → SNZ2
BK006947	12986	1				1		1						1		1	1			(T)11→14	intergenic (-110/-281)	SNO2 ← / → SNZ2
BK006947	189585			1	1		1	1	1						1		1	1	1	C→T	S512F (TCC→TTC)	SLA2 →
BK006947	779332				1															(A)24→28	intergenic (-594/-584)	AIF1 ← / → COS10
BK006948	92	1	1			1		1				1		1	1	1		1	1	Δ3 bp	intergenic (–/-491)	– / → YOL166W-A
BK006948	458209	1	1			1				1	1			1	1	1				G→C	intergenic (-389/+1271)	GYP1 ← / ← NRT1
BK006948	773264	1	1			1				1	1	1	1	1		1	1			G→C	E222Q (GAG→CAG)	MKK1 →

Table S3. ANOVA of fluorescence intensity ($Excitation_{485}$ & $Emission_{540}/Absorbance_{600}$) measured via a Nile Red spectrophotometric assay.

Fixed Effects	**F-value**	**Prob > F**
Gene of Interest (GOI)	$F_{5,223}$= 28.5	**<.0001**
Background Strain	$F_{1,223}$= 5.72	**0.02**
GOI*Background	$F_{5,223}$= 0.214	0.96

Dunnett's: Comparison relative to WT

Level	**Difference**	**p-Value**
cdc1-G149E	-16.55	0.00
gpm1-G95S	2.36	0.98
mkc7-G124C	-0.65	1.00
pib2Δ468	32.99	**<.0001**
tgl3-G289R	17.03	**<.0001**

Dunnett's: Comparison relative to BY4741

Level	**Difference**	**p-Value**
BY4742	6.26	**0.02**

Table S4. ANOVA of TAG abundance measured via LC-QTOF-MS.

Fixed Effects	**F-value**	**Prob > F**
Gene of Interest (GOI)	$F_{2,17}$= 20.2	**0.0001**
Background Strain	$F_{1,17}$= 5.91	**0.0318**
GOI*Background	$F_{2,17}$= 3.14	0.075

Dunnett's: Comparison relative to WT

Level	**Difference**	**p-Value**
pib2Δ468	1.86	**<.0001**
tgl3-G289R	1.31	**0.0017**

Dunnett's: Comparison relative to BY4741

Level	**Difference**	**p-Value**
BY4742	0.6	**0.03**

Table S5. Oligonucleotide sequences.

Name	Targeted gene	Oligonucleotide sequence
ADH7 gBlock	*ADH7*	CtttggtctcaccaaaacTTCAAATACCAGAAAATATTC CAAGTCCGCTAGCCGCTCCATTATTGTGTGGT GaTATTACAGTTTTCTCTCCACTACTAAGAAA TGGCTGTGGTCCAGGTAAGACCGCTCCATTAT TGTGTGGGttttagagagagacctttc
GPM1 gBlock	*GPM1*	CtttggtctcaccaaaacCAGATCCTGGAGATTGAACG AAAGACATTACGGTGACTTACAAaGTAAGGA CAAaGCTGAAACTTTGAAGAAGTTCGGTGAA GAAAAATTCAACACCTACAGAAGATGACTTA CAAGGTAAGGACAGttttagagagagacctttc
CDC1 gBlock	*CDC1*	CtttggtctcaccaaaacCAGTATTACTTGGATCCTGAT AGCAATTTTTTCCTAGGTGATCTATTTGATGG AGaGAGAAACTGGGATGATAAGCAATGGATT AAAGAGTACACACGATTCAACCACTAGGTGA TCTATTTGATGGGttttagagagagacctttc
IRS4 gBlock	*IRS4*	CtttggtctcaccaaaacGGTAAGCAACAGACACCGTC ATTTGAACCTGTTGAGCTGGTGGCCCTCTATA ACcGGTGATTCCGGAGtAATTAATACTTTACCC GAAGATGGATTGATATTGGGTGAGCTGGTGG CCCTCTATAAGttttagagagagacctttc
MKC7 gBlock	*MKC7*	CtttggtctcaccaaaacGGATACCACTtGCTCATCCTTT AAACAAGTAAACAAGGATGCTTTGGCTAGtGT tGTTGAGAGTGTGTTTACTGAAATTAGCTACG ACACCACTATAGTTACCAGCCAAGGATGCTTT GGCTAGCGGttttagagagagacctttc
PIB2 (Δ188) gBlock	*PIB2*	CtttggtctcaccaaaacCAAAGATCGAACCCTGAACT AGATCCCACCAATAGCGTGGTGGATGTTAGA GAATACAAAGAGCAAAACTGTGTTCAATGAG TTAGAAGATGATGCCGAGCGTGGTGGATGTT AGCAGGttttagagagagacctttc
PIB2new (Δ468) gBlock	*PIB2*	CtttggtctcaccaaaacTTCCGGACAGTAAGAGGAAT TCTTGTCGTTACTGCCATAAGCCTTTTACCCT GTaGtggcGGAAAgATCATTGCAGGCATTGCGG GGACATCTTTTGTCAAGACCAGCCATAAGCCT TTTACCGttttagagagagacctttc
UFD1 gBlock	*UFD1*	CtttggtctcaccaaaacTGCTACCCTATAGCTATGATG AATGATAGAATACGGAAGGATGACGCAAACT TCGaCGGGAAAATATTCCTGCCACCAAGTGCT CTAAGCAAACTATCCATGTGGAAGGATGACG CAAACTTGttttagagagagacctttc

Table S5 Continued		
Name	**Targeted gene**	**Oligonucleotide sequence**
IRS4m-F	*IRS4*	TCTATAACCGGTGATTCCGGcGT
IRS4w-F	*IRS4*	TCTATAACGGGTGATTCCGGcGC
ADH7-FP	*ADH7*	AGGTTGGATCCAAGTGCCACAC
ADH7-OUTF	*ADH7*	GGCGATCATGACGTTGATGTTGAA
CDC1-FP	*CDC1*	ATGAATCAGTGGTTGTGAAGAGA
CDC1-OUTF	*CDC1*	TACACACAGCAAGAAGAGCGACG
CDC1-RP	*CDC1*	ATTTCCAGCGTCTAATGAGGAGGA
GPM1-FP	*GPM1*	ACCCAGACGTCTTGTACACTTC
GPM1-OUTF	*GPM1*	AGTTAGACACGGTCAATCCGAATG
GPM1-RP	*GPM1*	TTACCACTCAACAAGTCCTTGGC
IRS4-FP	*IRS4*	TGCACATCATAGGCCTATCCAC
IRS4-RP	*IRS4*	CTTGTGTCTACCTTCGTGTAAAT
MKC7-FP	*MKC7*	ATTCGGTGGAATTAGATATTGGCA
MKC7-OUTF	*MKC7*	ACCAAGTAATGAAGACTATGTCA
MKC7-RP	*MKC7*	GCCAGGCAAACCAATAC
PIB2-OUTF	*PIB2*	CACAGTGACATCTCCCATCACAT
PIB-FP	*PIB2*	GTTGACAGAAAGCAGCAGATCCA
PIB-RP	*PIB2*	TTCGTTGTAGTCAGCTTGTCGACA
TGL3-FP1	*TGL3*	ATGAAGGAAACGGCGCAGGAATACAAG
TGL3-FP2	*TGL3*	CTGTTTCACCTTGGGGTGATTAGAGGT
TGL3-RP1	*TGL3*	CTCCGTTATGCTCGGATTGTTGGCATT
TGL3-RP2	*TGL3*	CCTACTCCGTCTTGCTCTTATTATGTC
UFD1-RPI	*UFD1*	CAGTAGATGAAATTTGCAGTAGA
UFD1-RPO	*UFD1*	CACATATCCTACAGGTGGAGCAA
UFD-FPO	*UFD1*	ATGTTTTCTGGCTTTAGTTCTT

CHAPTER 3. CHARCTERIZATION OF MORPHOLOGICAL AND CUTICULAR LIPID TRAITS AND THEIR RESPECTIVE ROLES IN MEDIATING RATES OF WATER LOSS FROM MAIZE SILKS

Bri Vidrine[1], Keting Chen[2], Nick Lauter[3], Basil J. Nikolau[1], & Marna D. Yandeau-Nelson[2]

[1]Iowa State University, Department of Biochemistry Biophysics & Molecular Biology, Ames, IA, and USA; [2]Iowa State University, Department of Genetics, Development and Cell Biology, Ames, IA, and USA; [3]Iowa State University, Department of Plant Pathology, Ames, IA, and USA

Key words: Surface lipids, Cuticular wax, Maize, Desiccation, Water loss, Silks

Abstract

Stigmatic maize silks are essential for pollen reception, hydration, tube growth and fertilization, however while maximization of surface area is advantageous for capturing wind carried pollen, these large exposed surfaces also increase susceptibility to water loss, especially under desiccating conditions. Extracellular cuticular surface lipids (SLs) that coat plant surfaces are thought to protect tissues against desiccation, and previous studies have demonstrated variation within the silk surface lipid metabolome. To investigate the protective capacity of extracellular surface lipids against water loss, this study implemented an experimental, in vitro system to evaluate the effects of both temperature and humidity on rates of water loss from excised silks. Emerged and husk-encased fractions of silks from two IBM recombinant inbred lines, MO145 and MO378) and the B73 and Mo17 parental lines were assessed for their rates of silk water loss under different temperature-humidity conditions in tandem with metabolomic profiling revealing that emerged silks from B73 and MO145 had the highest concentration of total surface lipids and were also had the lowest overall rates of water loss. Emerged silks were sampled over a four-year span to investigate the relative contributions from physical silk traits including, SL traits and non-lipid

traits. An optimized predictive model was constructed through a cross-validation approach and it identified two physical silk properties which predicted the most variation amongst individual silk samples, irrespective of genotype. Application of the model to new data generated the same general result: increased alkane accumulation was predictive of lower rates of water loss and higher initial silk water content (SWCi) predicted elevated rates of water loss. Together this study concludes that alkanes are likely protective against silk water loss across both high and low relative humidities, however, tissue hydration also seems to be important and the husk-encasement status (emerged or husk-encased) had a greater impact on rates of water loss in determining important in determining silk

Introduction

Terrestrial plants depend on water storage and maintenance for survival, which is exacerbated by their unrelenting exposure to the environment and unpredictably fluctuating weather conditions. The cuticle is the outermost physical barrier between aerial portions of a plant and the environment, protecting against both biotic (e.g. pathogen invasion and insect herbivory) and abiotic stresses (e.g. UV radiation, extreme temperature, and water limitation) (Barnes et al., 1996; Riederer & Schreiber, 2001; Shepherd & Wynne Griffiths, 2006). The hydrophobic cuticle is produced and secreted by epidermal cells and consists of a cutin polyester matrix that is infused with and coated by cuticular lipids, which are comprised of very-long chain fatty acid derivatives (e.g. alcohols, aldehydes, hydrocarbons, wax esters) (Li-Beisson et al., 2013). The composition of the cuticle varies among species, as well among organs and across tissue development within a plant (Jetter, Kunst, & Samuels, 2006; S. B. Lee & Suh, 2015).

Surface lipid abundance and composition is impacted by water stress and may act as an adaptive defense against desiccation in diverse plant species (Dylan K Kosma et al., 2009; Shepherd & Wynne Griffiths, 2006). For example, periodic water limitation stress (i.e. drought)

has been shown to increase total SL load in maize, rose, tree tobacco, soybean and sesame (Cameron et al., 2006; K. S. Kim et al., 2007; Dylan K Kosma et al., 2009; Premachandra et al., 1991). In Arabidopsis, water-stressed rosette leaves accumulated 75% higher levels of cuticular wax, comprised primarily of alkanes (Dylan K Kosma et al., 2009). In pine, non-irrigated saplings exhibited higher expression of specific cuticular wax biosynthesis genes (e.g. CER2, CER6, CER10 and LACS3) concomitant with increased cuticular waxes, as compared to irrigated saplings (Le Provost et al., 2013). Moreover, knockout studies of cuticle-related genes have suggested that surface lipids play a role in protecting against abiotic stress. For example, overexpression of *OsGL1-3* in rice increased tolerance to water deficit and this gene was found to be involved in cuticular wax biosynthesis suggesting that this biosynthetic function may contribute to tolerance to water deficit (Bourdenx et al., 2011; X. Zhou et al., 2015). Changes in total lipid accumulation have also been observed under drought stress in Arabidopsis, sesame, tree tobacco, and saltwater cress (Cameron et al., 2006; K. S. Kim et al., 2007; Dylan K Kosma et al., 2009; Xu et al., 2014).

The silks of maize are the stigmatic floral organs that facilitate pollination and are the conduits for fertilization of the ovule by pollen. The silk surface accumulates unusually high levels of aliphatic hydrocarbons with backbones ranging from 19-35 carbon atoms (Loneman et al., 2017; Perera et al., 2010). Their size, availability, and hydrocarbon-rich SL metabolomes make silks an excellent model for dissecting the hydrocarbon biosynthesis pathway and more generally, in studying stigmatic biology in monocots. The compositional requirements of The composition of the silk SL metabolome may be important to both protect silks against water loss and also to facilitate pollination. Environmental exposure is mediated by early, husk-encased growth which requires physical separation of individual silks for proper development and is

thought to be mediated by the cuticle, as specific cuticle-deficient mutants exhibit a fused organ phenotype (Yeats & Rose, 2013). Inbred lines B73 and Mo17 were shown by to preferentially increase surface lipid deposition (primarily hydrocarbons) on emerged silk surfaces by 2 to 4-fold relative to their husk-encased counterparts (Loneman et al., 2017; Perera et al., 2010). Therefore, this surface lipid accumulation may play a protective role and variation in the surface lipid composition could correspond to an advantage in select environments.

In this study, the impacts of cuticular lipid abundance and composition, as well as other traits, on the susceptibility of maize silks to desiccation were evaluated for excised silks from four genotypes (inbred lines B73 and Mo17, and recombinant inbred lines MO145 and MO378) that were exposed to different temperature and relative humidity conditions. Rates of water loss were compared among genotypes, between emerged and husk-encased and across environments. Statistical models were generated and tested that included both surface lipid and non-surface lipid traits (e.g. silk water content). A four-year study is presented that demonstrates that surface lipids, particularly alkanes, and silk water content are predictive of rates of water loss from silks.

Results

Method for measuring rates of water loss from excised silk tissue

Rates of water loss were determined from silks that had emerged from the encasing husk-leaves into the external environment and were harvested from field-grown MO378 plants. The sites of silk excision are likely a primary site of water loss, in addition to water loss that occurs through the silk cuticle. To assess the impact of cutting the silks, we assessed the extent of water loss from excised silks under desiccating conditions (34°C and 15.8% relative humidity). First, a 3 cm segment was excised from silks, generating two freshly cut ends per silk fragment. Excised silks were either maintained as 3 cm segments (Control) or bisected into two 1.5 cm lengths, generating a fresh pair of sheared ends, either at the time of treatment (0 hr) or after 30 min (0.5

hr time point) of desiccating treatment. Silk mass was recorded at 30 min intervals over a 4-hour period. Samples cut at initiation of the desiccating conditions experienced significantly higher rates of water loss during the first half hour (0-0.5 hr interval) as compared to the control samples (**Fig. 1**). In subsequent intervals however, rates of water loss did not differ between the silks in the control treatment as compared to those silks that were bisected at time 0. Similarly, bisection of silks after 30 min in desiccating conditions (t=0.5 hr) resulted in an increased rate of water loss during the 0.5-1hr interval as compared to the control. Similar results were observed for this experiment conducted on silks from inbred lines B73 and B104 under ambient conditions (data not shown). Because rates of water loss equilibrate within 30 min after cutting the silks, subsequent experiments measure silk mass starting 1 hr after silks are cut into 3 cm segments and placed into treatments of different temperature and relative humidity combinations.

Rates of water loss differ between husk-encased and emerged silks and are impacted by genotype, environment and GxE interactions

Cuticular hydrocarbon accumulation has previously been shown to be 2- to 4-fold higher on emerged portions as compared to the husk encased portions of silks (Loneman et al., 2017; Perera et al., 2010). Thus, we hypothesize that the increased cuticular lipid accumulation in emerged silks will results in decreased rates of water loss as compared to husk-encased silks. Rates of water loss were measured from emerged and husk-encased silks from four genotypes: two parental inbred lines B73 and Mo17, and two Intermated B73 x Mo17 Recombinant Inbred Lines (IBMRILs), MO145 and MO378. Both IBMRILs accumulate lower total surface lipid loads than the parental inbred lines. Moreover, relative surface lipid compositions differ between the two IBMRILs. Like B73, MO145 accumulates higher concentrations of hydrocarbons (alkanes and alkenes), whereas MO378, like Mo17, accumulates higher concentrations of fatty acids (**Fig. S2**). The diverse surface lipid compositions amongst the four genotypes permits

examination of the role of hydrocarbons and fatty acid composition and the role of total surface lipid abundance on rates of water loss from silks.

Because both temperature and humidity can impact the rate of water loss from plants, emerged and husk-encased silks were exposed to six controlled combinations of temperature (15°C, 25°C and 35°C) and relative humidity (15% and 85%) for four hours, as described in the Materials and Methods. Within each glass chamber, environmental data loggers recorded the actual temperature and relative humidity, which were used to calculate the vapor pressure deficit (VPD) to estimate the desiccating potential of each treatment. Rates of water loss differed among genotypes, according to husk encasement status and across-temperature-humidity treatments (Figure 2; ANOVA analysis in Table 1). For example, rates of water loss were elevated in emerged silks relative to husk-encased silks at 35°C, with increases ranging among inbred lines from 1.2 to 1.6-fold and 1.3 to1.7-fold under the 15% and 85% relative humidity treatments, respectively. These differences in rates of water loss were not observed under the 25°C treatment for the B73 and MO378 genotypes. The estimation of least squares (LS) means allowed for the assessment of the relative contributions of each effect (i.e. genotype, husk-encasement, temperature and humidity) to rates of water loss. Regardless of temperature-humidity treatment or silk husk-encasement status, B73 and MO145 silks exhibited the lowest rates of water loss (approximately 0.46 mL/g dry silk/hr) and MO378 exhibited the highest (0.67 mL/g dry silk/hr; **Table S2**). Husk-encasement status also impacted water-loss rates. Emerged silks exhibited ~35% higher rates of water loss compared to husk-encased silks according to their least square means, which contradicts the hypothesis that emerged silks would be more resistant to desiccation than husk-encased silks. Unsurprisingly, environmental factors had a significant impact on rates of water loss, with temperature and relative humidity explaining ~24% and 33%

of the observed variance, respectively (Table 1). The 35°C treatment induced ~1.9-fold higher rates of water loss relative to the 25°C treatment and ~2.1-fold higher rates under the 15% relative humidity to the 85% relative humidity treatment.

In addition to the impact of individual factors on rates of water loss, several interactions also had significant impacts, including temperature x relative humidity, genotype x relative humidity, husk-encasement status x temperature, and genotype x husk-encasement status. The interaction between genotype and husk-encasement status (**Table 1**) is exemplified by the observation that emerged silks from inbred Mo17 were less resistant to water loss (i.e. higher water loss rates) than emerged silks from inbred B73, whereas rates of water loss for husk-encased silks did not statistically differ between these genotypes. The significant interaction between temperature and relative humidity suggests that the 35°C with 85% relative humidity, and the 25°C with 15% relative humidity treatments did not cause statistically different rates of water loss despite the 2-fold difference between their respective vapor pressure deficits (**Fig. 2**). The observed interaction between husk-encasement status and temperature (Table 1) is exemplified by the observation that rates of water loss from emerged silks is 1.4-fold higher than husk-encased silks at 35°C, whereas no difference is observed relative to husk-encasement status under the 25°C treatment. Collectively, these data suggest that emerged silks are more susceptible to water loss at high temperatures than their husk-encased counterparts. Pronounced differences in rates of water loss amongst genotypes under the 15% relative humidity treatment were observed, however under the 85% relative humidity treatment, genotypes could not be distinguished from one another (**Table S1**). This suggests that a genetically-determined property is affecting performance selectively between relative humidities, which is supported by the

observation that a genotype x relative humidity interaction contributes modestly (~2%) to the variance observed in rates of water loss (Table 2).

Surface lipid profiling shows that emerged silks from B73 and MO145 accumulate higher concentrations of total surface lipids

The composition of the surface lipid metabolome was assessed from companion samples of emerged or husk-encased silks used to assess rates of water loss (Fig. 2). A representative sub-sample of silks per biological replicate pool was assessed for surface lipids via GC-MS without subjection to the temperature and relative humidity treatments. Total accumulation of surface lipids was highest for emerged silks from the B73 and MO145 genotypes, exhibiting 1.7-fold and 2-fold higher concentrations relative to their husk-encased fractions, respectively (**Fig. 3A**). In contrast, husk-encasement status did not impact total surface lipid accumulation for MO378 and Mo17 silks. Across inbred lines, emerged silks from B73 and MO145 exhibited approximately 2-fold higher SL load as compared to MO378 and Mo17. Surface lipid composition, particularly the relative abundance of hydrocarbons, did not significantly change relative to husk-encasement status. However, minor variations, up to 5%, were observed amongst genotypes, by up to 5% (**Fig. 3B**). Similarly, the relative abundance of alkenes relative to total hydrocarbons was similar between emerged and husk-encased silks within each genotype. However, the relative abundance of alkenes varied among inbred lines, with a 10% difference observed for emerged silks from inbred lines B73 and Mo17 (**Fig. 3C**).

Both non-SL and SL traits impact rates of water loss from silks

Genotype and husk-encasement were both shown to impact rates of water loss (**Table 1**). Differences in rates of water loss may be captured by both factors; however, they may mask the underlying physical trait(s) that govern this response. Herein, we consider the contributions of both surface lipid traits and non-lipid silk traits to observed rates of water loss (**Fig S3**). Surface

lipid traits include the abundances of alkanes, alkenes, and fatty acids as well as relative abundance traits including percent hydrocarbons of total SLs, and percent alkenes of total hydrocarbons (**Fig. 3**). In addition to these lipid traits, we also consider both silk water content and silk length, relative to silk dry weight. These traits can be considered a proxy for gross morphological traits that may affect surface area. Silk water content at the time of excision from the ear was similar amongst all genotypes. Similarly, silk water content did not differ between emerged and husk-encased silks, except for Mo17, for which water content was slightly lower in emerged as compared to husk-encased silks (**Fig. S3**). The ratio of silk length relative to dry weight provides insight into, however differed relative to both husk-encasement status and genotype. For example, silk length relative to dry weight was lower (i.e. higher biomass for a fixed length) in both emerged and husk-encased silks for Mo17 as compared to MO378. Unlike for Mo17, this ratio varied relative to husk-encasement status in MO378, with silk length relative to dry weight being lower in emerged silks (which would correspond to longer silks for a given dry weight).

To test the impacts of both SL traits (**Fig. 3**) and the non-SL silk traits (**Fig S3**), as well as environmental parameters (i.e. temperature and relative humidity) on rates of water loss, each of these trains was treated as a continuous variable in an ANOVA model. Both SL- and non-SL silk traits contribute to rates of water loss. For SL silk traits, abundances of both alkanes and fatty acids, as well as percent hydrocarbons of total SLs, in combination explain 4% of the observed variance in rates of water loss (**Table S1**). The non-SL traits, silk length relative to dry weight and silk water content, explain 7.5% and 5.3% of the observed variance, respectively. In summary, all individual physical silk traits were found to be significant factors within this model, except for alkene abundance and % alkenes of total hydrocarbons. The significant factors

cumulatively explained 79.5% of the variance, which was comparable to results obtained using categorical variables, which explained 76.8% (Table 1). These results demonstrate that these SL- and non-SL physical silk traits can explain a similar amount of the variance compared to a statistical model that includes only genotype and husk-encasement status.

Rates of water loss amongst the four genotypes are similar across a 4-year study

To further explore the potential relationships between rates of water loss and surface lipid metabolome composition, we assessed the four genotypes (B73, MO145, MO378, Mo17) over the course of four successive growing years (2015-2018). Because lipid variation amongst these four genotypes was larger in emerged silks relative to husk-encased silks (**Fig. 3**), this experiment focused on emerged silks only. In addition, because rates of water loss were most impacted at higher temperature, excised silks were exposed to either 15% or 85% relative humidity at a single temperature (35°C).

Emerged silks were assessed for both rates of water loss and surface lipid composition in each of the four years (**Fig. 4**). Comparisons amongst genotypes within each year at either 15% or 85% relative humidity only identified significant differences in rates of water loss within the 2018 growing year, presumably due to the lower number of replicates within each year. In 2018, silks from MO378 exposed to 15% relative humidity exhibited ~1.3-fold higher rates of water loss compared to the other inbred lines. At 85% relative humidity, MO378 exhibited 1.4-fold higher rates of water loss relative to B73 (**Fig. 4A**).

Lipid metabolomes amongst the four genotypes are similar across a 4-year study

Major differences in surface lipid composition amongst these four genotypes were observed across growing seasons (**Fig. 4B**). MO145 accumulated the highest average abundance of total surface lipids, whereas MO378 accumulated significantly less than the other three genotypes. Surface lipids accumulated to significantly lower levels in 2017 relative to the other

three years, regardless of genotype. Of the four genotypes, Mo17 exhibited the lowest relative abundance of hydrocarbons (out of total SLs) than the other three genotypes: Mo17 (78%) vs B73 (92%), MO145 (91%), and MO378 (87%). Higher fatty acid accumulation accounted for this difference. Similarities between B73 and MO145 were observed for their hydrocarbon chain-length preferences; both produced appreciable quantities of tricosane (C23) and pentacosane (C25). This pattern differed as compared to the inbreds Mo17 and MO378, which accumulated very few hydrocarbons of these acyl chain lengths and instead produced higher amounts of heptacosane (C27). The consistency of metabolite traits observed across years, particularly when considering trends for individual metabolites, provides evidence that the metabolome is under robust genetic control.

Total surface lipid accumulation was observed to be inversely related to rates of water loss when silks were subjected to 15% relative humidity. MO378 accumulated ~50% lower total surface lipids as compared to the other inbred lines and exhibited the highest rate of water loss. Rates of water loss are significantly impacted by relative humidity, genotype and growing year within the 4-year study (ANOVA Total R^2=X, p-value<0.5; **Table S3**). In three of the four years, rates of water loss did not statistically differ among the genotypes (**Figure 4)**. However, genotype was found to significantly contribute to rates of water loss when all years were considered together. As expected, relative humidity again was a significant factor and rates of water loss were found to be ~1.8-fold higher under the 15% relative humidity relative to 85% relative humidity treatment according to least square means. The growing year also contributed to the observed variance in rates of water loss.

SL composition predicts lower rates of water loss

To assess the impacts of physical silk traits on rates of water loss across the 4-year study, an optimized statistical model was generated, consisting of the most influential continuous

factors. The same physical silk traits included in the ANOVA reported in Table S1 (the SL traits, alkanes, alkenes, fatty acids, percent hydrocarbons of total SLs, percent alkenes of total hydrocarbons; and the non-lipid silk traits, silk water content, and silk length/DW). These traits were considered alongside environmental parameters, including growing year and observed relative humidity. To identify traits within the model that are consistently selected as contributing factors to rates of water loss, a step-wise regression and cross-validation approach was used in which the 4-year data set was selected and tested through random training and validation subsets of the dataset (**Fig. 4**). Iterative testing identified terms that were repeatedly selected, and thus likely important in the model. Ten iterations of 10-fold cross-validation (100 total repetitions) identified 11 terms as important in over 90 repetitions (**Fig. 5)**, including three lipid traits (alkanes, alkenes and percent alkenes of total hydrocarbons) and interaction of these traits with growing year, two non-lipid silk traits (silk water content and silk length/DW), and two environmental parameters (growing year, relative humidity, and an interaction between relative humidity and growing year).

Subsequent multi-factor ANOVA with the 11 terms selected by cross-validation identified seven terms that collectively explained 79% of the observed variance in rates of water loss (**Table S4**). According to this ANOVA analysis, the other four terms did not meet the significance threshold ($p<0.05$) and were therefore removed. The seven remaining terms were used to construct the final, optimized model, which was subsequently assessed through an alternative, Bayesian hypothesis testing approach. Terms from the 7-factor full model were removed individually, to generate seven reduced models, each with six factors. These reduced models were then compared to the full model to assess impact on overall model performance and resulting Bayes Factors were ranked to determine the relative impact of each factor (Kass &

Raftery, 1995). In addition, regression coefficients were calculated for each of the seven terms to indicate either a positive or negative relationship with rates of water loss (**Table 2**). Relative humidity was the most important predictor of rates of water loss, with low relative humidity inducing higher rates of water loss (**Figure 4A)**. The second most important term, alkane abundance (BF=46.9), was negatively associated with rates of water loss. Therefore, silks with higher alkane accumulation tended to lose water less rapidly than silks with low accumulation, suggesting that alkanes may play a protective role against water loss. Thirdly, initial silk water content was a positive predictor of rates of water loss, silks that were more hydrated (relative to silk dry weight) initially, tended to lose water more quickly. Finally, growing year was negatively associated with rates of water loss.

The robustness of this optimized model explaining the contributing factors to rates of water loss (**Table 2**) was tested using a separate validation dataset consisting of both silk surface lipid composition data and rates of water loss measured at 15% and 85% relative humidity and 35°C, collected from all four genotypes during growing year 2016. Notably, these data were not included in the 4-year dataset. Rates of water loss within the validating dataset revealed genotype-specific effects (**Fig. S4**). Mo17 exhibited 1.4-fold and 1.8-fold higher rates of water loss than B73 under the 15% and 85% relative humidity treatments, respectively. In contrast, MO378 only exhibited statistically higher rates of water loss (i.e. 1.6-fold) as compared to B73 under the dry, 15% relative humidity treatment. Surface lipid profiling also revealed distinct differences in metabolome composition across the four genotypes. Total surface lipids were ~2-fold higher in MO145 relative to MO378 (**Fig. S5A**), whereas B73 and Mo17 were not statistically different from any of the genotypes in total surface lipid accumulation. The relative abundance of hydrocarbons relative to total SLs was ~4% higher in B73 and MO145 than Mo17

and MO378 (**Fig. S5B**). Thus, Mo17 and MO378 accumulate higher levels of fatty acids. Out of all hydrocarbons, alkene relative composition was ~12% higher in B73 relative to Mo17 and the MO378 and MO145 genotypes fell between the B73 and Mo17 levels, with ~58% alkenes out of total hydrocarbons (**Fig. S5C**).

To test whether the statistical model generated from the 4-year dataset was predictive rather than just estimative, the model was applied to the new 2016 dataset. Because only one year of data was represented in this dataset, the terms that included year in the optimized model were not considered. The five-term model included the terms, observed relative humidity, alkane abundance, alkene composition relative to hydrocarbons, initial silk water content, and silk length per dry weight. This model explained 63% of the observed variance in rates of water loss, (**Table 3**), with most of this variance contributed by the relative humidity factor (partial R^2=0.52). Importantly, the two physical silk traits that were found to be important according to Bayes factor analysis, alkane abundance and initial silk water content (**Table 2**), were also significant in this validating model. Overall this supports the conclusion that these physical silk traits, initial silk water content and alkane abundance, are predictive of rates of water loss at 35°C.

Comparison of several silk morphological traits across inbred lines

In this study, we have shown that initial silk water content impacts rate of water loss, suggesting that silk morphology is important. Emerged silks were found to lose water more rapidly than their husk-encased counterparts despite increased surface lipid deposition on emerged silk surfaces (**Fig. 2** & **3**). Silk surface area, which could be influence by trichome size and number, likely impact rates of water loss. Therefore, morphological features, including water content, cross-sectional area of the silk, and trichome density, were surveyed across the four inbred lines (**Figure S6)**. These three measures were more similar amongst genotypes for husk-

encased silks relative to emerged silks. Mo17 had the largest perimeter, an average of 1.49 cm at the base and 1.4 cm at the top of the emerged section, respectively. Mo17 also had the largest cross-sectional area out of the four genotypes at the distal end of emerged silk segments. Despite having a larger perimeter and area for emerged silks, Mo17 also had relatively short silk lengths per dry silk weight (~90m/g vs ~105-130m/g; **Table S5**) indicating that Mo17 emerged silks were wider and shorter as compared to B73 and MO145 silks. Trichome surface area was also estimated on emerged portions of silks (**Figure S7**). Trichome surface area was over 30% larger in Mo17 silks relative to the other inbred genotypes and ~2-fold higher than trichomes from MO378 silks. Trichome number per silk length was ~25% higher for Mo17 and MO378 as compared to B73 and MO145 however this difference was only present in emerged silk tissue (**Figure S6**). Because the genotypes that tended to exhibit higher rates of silk water loss (i.e. MO378 and Mo17) also had more trichomes on their surfaces, trichome number may also influence rates of water loss.

Discussion

This study considered the agronomically important trait, rate of water loss from silks, in the context of changing environmental conditions. Through model optimization and subsequent validation, increased cuticular alkane abundances were determined to predict lower rates of water loss. The alkane content within the cuticle could thus play a protective role under hot temperatures (35°C). In contrast, alkenes were not found to explain a portion of the variance in observed rates of water loss and were thus not included in the optimized model. However, due to the correlation between alkanes and alkenes, alkenes levels may also be important. The potentially protective role of alkanes has been suggested by studies in other plants, such as saltwater cress and soybean, in which exposure to desiccation stress induced cuticular lipid accumulation and conditioned the plant to be more drought resistant (K. Kim, Su, Park, Kim, &

Jenks, 2007; Xu et al., 2014; X. Zhou et al., 2014). An alternative explanation for the predictive power of alkanes is that their levels could correspond to other traits that are important that were not included in the initial model (e.g. cuticle thickness, cutin composition, etc.).

Rates of water loss were higher in emerged silks, even with a larger load of surface lipids, suggesting that high surface lipid abundance is not the only cuticular trait that impacts cuticle permeability. are less important than husk-encasement status. Indeed, it has previously been shown that increased cuticular thickness is not always positively correlated with cuticle impermeability (Riederer & Schreiber, 2001). Further, the silk cuticle can be discontinuous (Heslop-Harrison, Heslop-Harrison, & Reger, 1985), which has led to the hypothesis that such cuticular discontinuities on emerged silks provide entry-points for pollen tubes during pollination. However, these same cuticular discontinuities may also facilitate increased rates of water loss as compared to husk-encased silks (Heslop-Harrison et al., 1985). This study also identifies silk water content as an important factor in determining the rates of water loss from silks. Although higher initial water contents were predictive of increased rates of water loss, the underlying reason for this is unclear. One possibility is that a higher ratio of water relative to silk dry weight reflects a larger surface area from which water can be lost.

Understanding the protective role of surface lipids relative to silk water loss will require further study. Alkanes were implicated for their role in moderating rates of water loss, however, they explained less of the observed variance than does genotype (1.4% vs 7%; **Tables S1** & **1**). Therefore, this study likely does not have enough power to resolve potential interactions between environmental conditions and metabolite traits. Alternatively, other silk traits may explain some of these genotypic differences and if these traits were not included in the model these factors may still need to be considered to explain the variance observed. that were not incorporated into

this model. Ideally, an alkane-deficient mutant could be tested to directly assess whether the mutant exhibits higher rates of water loss, as has been tested in previous studies of rice and Arabidopsis among others, and this could be expanded to different environmental conditions (Lu et al., 2012; X. Zhou et al., 2015). However, most maize mutants were identified at the seedling stage and as of now, no mutants are known to express significant reductions in alkane content on silks. Therefore, alternative tissue types or organisms should be considered to address the role of surface lipids across environments. Arabidopsis is a good candidate because its stems accumulate high alkane concentrations and mutants are available with known impairments in alkane production in the stems (Aarts, Keijzer, Stiekema, & Pereira, 1995; Bourdenx et al., 2011; Matthew A. Jenks, Tuttle, Eigenbrode, & Feldmann, 1995). Further associative analysis of the role of lipids in maize silks could be expanded to a more diverse population. For instance, Dennison et al. characterized 32 maize inbred lines that varied ~10-fold variation in total surface lipid accumulation (Dennison et al., 2019). This would improve the variation available for analysis, as opposed to ~2-fold genotypic variation observed in this study. Increasing the number of genotypes examined may also reduce the likelihood of confounding correlations between physical silk traits, both non-lipid and lipid. Focusing future studies on either environmental effects in alkane mutants or on a diversified maize silk panel will serve to maximize experimental power so potential interactions between surface lipids and the environment can be resolved and to dissect the role of surface lipids in protecting the unique and agronomically relevant silk (Anderson, Lauer, Schoper, & Shibles, 2004). Collectively, his study sheds light on important factors that impact water loss through the cuticle and suggests the importance of other silk traits in modulating plant water content under adverse environmental conditions.

Methods

Plant growth and silk collection

Inbred lines B73 and Mo17, and two recombinant inbreds MO145 and MO378, derived from the Intermated B73 x Mo17 Syn4 recombinant inbred line collection (M. Lee et al., 2002) were grown in four successive summers (2015-2018) at the Iowa State University Agricultural Engineering and Agronomy Research Farm (Boone, Iowa). Plants were grown under standard cultivation practices with no supplemental irrigation. The topmost ear shoot on each plant was covered prior to silk emergence to prevent pollination. Immature ears were collected three days after silks had first emerged from the encasing husk leaves, as described in previous studies (Loneman et al., 2017; Perera et al., 2010). Three ears were collected per biological replicate pool, with collection occurring between 9am and 1pm, local time, and transported to the laboratory at ambient temperature.

Tissue Processing

Per genotype, three ears were sliced longitudinally into halves, and three halves were pooled and mixed together to comprise a single biological replicate pool. Husks were removed and to maximize uniformity of the sample, the basal and tip portions of the cob were removed. Sub-samples from the emerged portions of the silks were obtained by cutting the silks 3 cm distal to the point of silk emergence and using the silk section that included the silk tip. Husk-encased silks were cut 3 cm proximal to the point of emergence and this portion was discarded; the subsequent 3 cm length of silk was collected for analysis (**Fig S1**). Emerged silk segments from three ears (biological replicates) were pooled together to create a biological replicate pool. Husk-encased silks were pooled together in the same fashion. Each biological replicate pool was split amongst environmental treatments, with one portion retained for profiling surface lipid

metabolites. In 2015, emerged and husk-encased silks were evaluated and in subsequent years only emerged silks were evaluated.

Environmental Treatments

Controlled temperature and relative humidity treatments were applied using saturated salt solutions within closed, glass desiccators as described previously (Fonseca & Westgate, 2005), with modifications. Saturated solutions of either lithium chloride or sodium carbonate were added to the bottom of each glass desiccator (PYREX 2.4L Small Knob Top Desiccator, Product No. 3081-150) to generate relative humidity treatments of 15% and 85%, respectively. To maintain homogeneous humidity within each desiccator, computer fans (Long Shen Xin, Product No. 1160702) powered by four AA-batteries were placed just above the saturated salt solutions. A data logger (Onset Computer Corporation, Bourne MA, HOBO U12-012) was mounted above the salt solution. Finally, an internal rack system made from wooden skewers and holding eight 3 cm weigh boats, was placed above the data logger. Each glass desiccator was then placed at constant temperature within incubators programmed to 25, 30, or 35°C. Relative humidity and temperature observations recorded by the data logger at 1-min intervals were used to calculate the actual vapor pressure (AVP) and the saturated vapor pressure (SVP), which were in turn were used to calculate the environmental vapor pressure deficit (VPD). VPD is a measure of the desiccating potential of the air and is commonly used to monitor environmental stress from hot and/or dry conditions (high VPD) and conditions that are excessively cold and/or damp (low VPD). VPD was calculated in kilopascals as follows (ASHRAE Research, 1993):

$$VPD\ (\text{Pa}) = SVP\ (\text{Pa}) - AVP\ (\text{Pa}) = \left[610.7 * 10^{7.5*Temp/(237.3+Tem\ )}\right] - \left[SVP * \frac{RH}{100}\right]$$

Water loss measurements

The experiment was organized in a random complete block design. Silk samples were randomly assigned to sample dishes within each treatment and silk fresh masses were recorded prior to initiation of treatments. Subsequently, silk masses were measured at 0, 1, 2, 3, and 4 hrs after initiation of temperature and relative humidity treatments, and silk water content was calculated as follows:

$$Silk\ water\ content\ \left(\frac{mL}{dry\ silk\ (g)}\right) = \frac{Silk\ mass - Dry\ silk\ mass}{Dry\ silk\ mass} \quad [1]$$

Missing dry weights (~2% of all collected) were imputed by multiplying sample fresh weight with the averaged dry weight per fresh weight ratio that was calculated from other samples from the same silk pool. Rates of silk water loss (Equation [2]) were calculated from the linear slopes of water content across four consecutive time points (1, 2, 3, and 4 hr).

$$Rate\ of\ water\ loss\ \left(\frac{\frac{mL}{dry\ silk\ (g)}}{hr}\right) = -slope\left(\frac{Silk\ water\ content\ (mL/g)}{time\ (hr)}\right) \quad [2]$$

After data collection, silk samples were dried at 35ºC for three days and both the number and dry weight of the silks were recorded for each sample. The silk length per dry weight was calculated by multiplying the number of silk fragments by 3 cm, and water content was determined from the sample fresh weight; both were normalized by silk dry weight.

Surface lipid extraction and gas chromatography-mass spectrometry analysis

Surface lipids were extracted from ~1 g of fresh silk by immersion for 4 min in hexanes:diethyl ether (9:1) supplemented with the standards eicosane and nonadecanoic acid at a concentration of 1 µg/mL of extraction solvent, as previously described (Loneman et al., 2017). Extracts were concentrated under a stream of N_2 gas in a N-EVAP nitrogen evaporator (Organomation Associates, Inc., MA). Prior to analysis via gas chromatography, extracts were

chemically derivatized via (Kates, 2010), with the following modifications. Briefly, the lipid extract was incubated in 2 ml of methanolic HCl (1N) for 1 hr at 85˚C. After incubation, 2 ml of 0.9% NaCl was added and the lipids were sequentially extracted with two aliquots of 0.5 ml hexanes. The non-polar hexanes layer was recovered after centrifugation for 3 min at 700 *x g*. Derivatized samples were dried under a stream of N_2 gas and reconstituted in hexanes prior to analysis via gas chromatography.

Gas chromatography coupled to a flame ionizing detector (GC-FID) was conducted using an Agilent Technologies series 6890 gas chromatograph equipped with a DB-1 cross-linked 100% Dimethylpolysiloxane column (30m in length; 0.25-mm inner diameter), and an Agilent Technologies model 5873 mass detector. Gas chromatography-mass spectrometry (GC-MS) was conducted using a Agilent Technologies series 6890 gas chromatograph equipped with an HP-5MS cross-linked (5%) diphenyl (95%) dimethyl polysiloxane column (30m in length; 0.25-mm inner diameter), or with a DB-1 cross-linked 100% Dimethylpolysiloxane column for peak identification, with helium as the carrier gas. Metabolites were detected via an Agilent Technologies mass spectrometer (Model 5973). For both GC-MS and GC-FID analyses, extracts (1 µl) were introduced to the column via a splitless injection, and the GC oven was initially at 80˚C, and the temperature was programed to increase by 15˚C/min to 220˚C, then by 7.5/min to 308˚C, 20˚C/min to 340˚C and finally maintained at this temperature for 2.33 min. GC-MS analysis was employed to determine the identity of metabolites and then GC-FID equipped with the same column was employed to quantify identified metabolites.

Retention indices of lipid metabolites were calculated based on calibration with straight chain hydrocarbons (C7-C40; Sigma Aldrich). Quantification of peak areas from GC-MS data was determined using the AMDIS software package and GC-FID data was analyzed using

ChemStation (Stein, 1999; Agilent). The NIST Mass Spectral library (http://webbook.nist.gov/chemistry/) was used for compound identification from mass spectral data. Initial concentrations of the internal standard, eicosane (C20 alkane), and dry sample weights were used to calculate metabolite abundances (μmol/g dry silk) from peak areas.

Morphological measurements of silks

For each of the four genotypes evaluated in 2016, three emerged silks were randomly selected from each ear and mounted in immersion oil to prevent water loss. Trichome length (mm) and trichome density (trichome #/mm of silk) were measured from silks imaged at 20X magnification on an EVOS™ fl Digital Inverted Microscope or at 2X magnification on a Nikon SMZ-U Stereoscopic Zoom Microscope, respectively. Trichome lengths and trichome number were determined from image analyses using Adobe Photoshop CC (V14.0). Images of trichomes were selected with the polygonal lasso tool and magic wand tool (default tolerance=42). The base of the trichome was cut off with a straight line using the polygonal lasso tool. Perimeter and area of the silk were recorded in pixels and the ruler tool was used to obtain a measurement of the linear base. Pixels were converted using a scale bar.

For imaging of cross-sectioned silks, silks were bundled in parafilm and sectioned into four portions. Silk ends were stained with aniline blue, which stains cell walls and vasculature. Immediately after staining, cross-sectioned silks were imaged at 10X magnification on a Nikon SMZ-U Stereoscopic Zoom Microscope. Trichome area (mm^2) and perimeter (mm) were measured from the images using Adobe Photoshop CC (V14.0). Specifically, each cross-sectioned silk was outlined and the area and perimeter was calculated from pixel areas normalized to the pixel area of a 1.5 mm pin head. A total of five silks were imaged for each genotype.

Statistical Analysis

Tukey's HSD and pairwise Student's T-tests were conducted in JMP Pro 14 (SAS Institute Inc., 2019). Analysis of variance was conducted using the aov() function in the R/stat package (R Core Team, 2019).

Important predictors of water loss rates were selected by forward and backward stepwise regression with the stepAIC() function in the R/MASS package (Venables, W.N., Ripley, 2002). To test predictive performance, we performed 10-fold cross-validation a total of ten times to generate 100 evaluated models.

Bayesian regression modeling was performed using the R/BayesFactor package (Morey & Rouder, 2015) to generate full as well as reduced models wherein a single factor was removed. Bayesian model testing was conducted using the function compare() in R/BayesFactor, and generated a Bayes factor for each factor tested in the model. Bayes factors >10 were considered to greatly impact the response viable (i.e. rate of water loss).

Acknowledgements

This work was facilitated by the W. M. Keck Metabolomics Research Laboratory at Iowa State University.

Figures

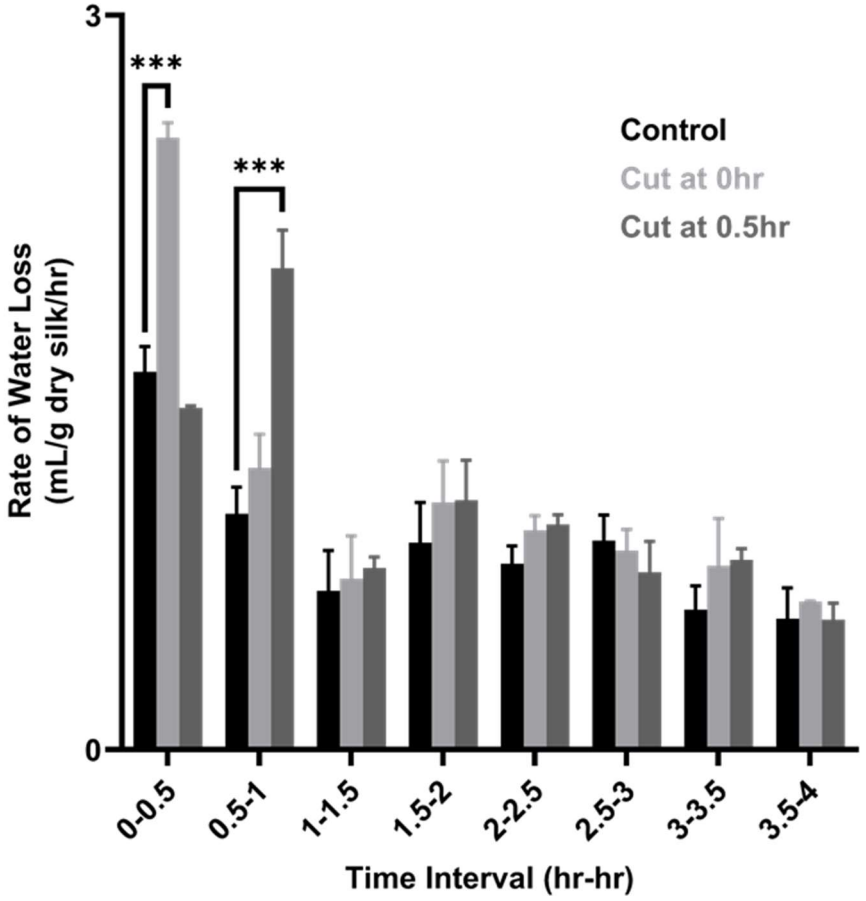


Figure 1 | The effect of silk bisection on the rate of water loss. Silk water content was measured from serial weighings of emerged silks from recombinant inbred line, MO378 exposed to 34°C and 16% relative humidity. Silks were either maintained as 3 cm segments (Control) or bisected into 1.5 cm sections at time 0 or 0.5 hr. The average rates of silk water loss are presented across 0.5 hr time intervals. Asterisks denote statistically significant differences between cut treatments relative to the control at each time interval (Dunnett's t-tests; *, $p<0.1$; **, $p<0.05$; and ***, $p<0.01$). Mean ± SE; 3 biological replicate pools, with two technical replicates measured for each. Chamber conditions were recorded to be 34.0 ± 1.7°C and 15.8 ± 2.0% relative humidity.

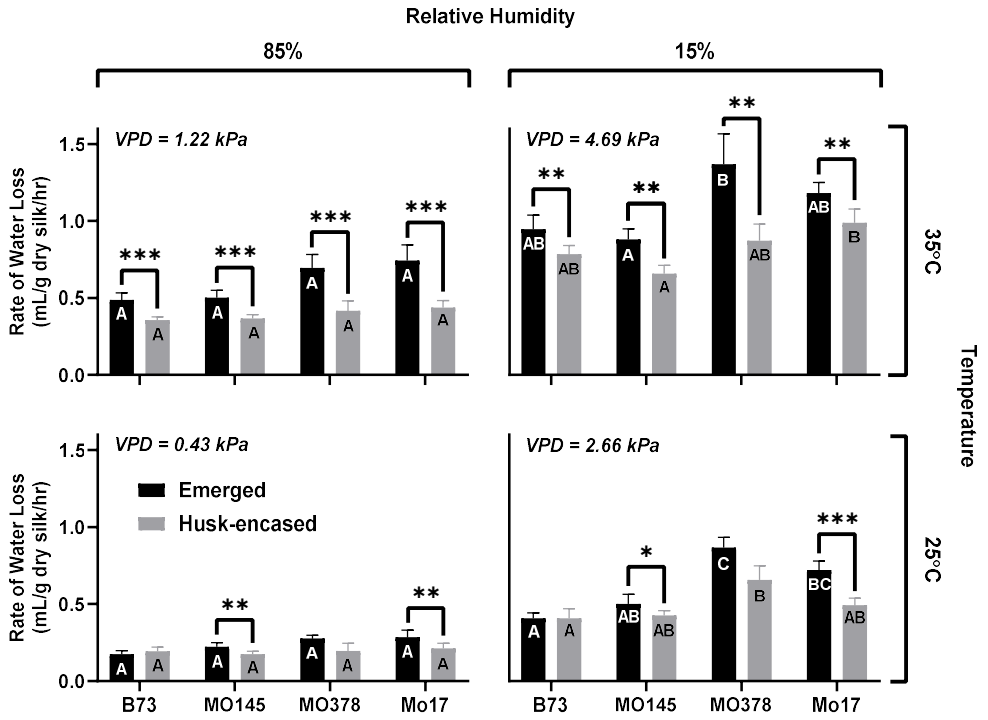


Figure 2 | Rates of water loss from emerged and husk-encased silks exposed to combinations of humidity and temperature treatments. Rates of water loss from emerged (black) and husk-encased (grey) silks. Average vapor pressure deficit (kPa) for each treatment is presented in the upper left of each graph. Significant differences in water loss rates between husk-encasement statuses (i.e. emerged vs. husk-encased) for each genotype were determined for each combination of temperature and relative humidity (t-tests; *, $p<0.1$; **, $p<0.05$; and ***, $p<0.01$). Mean ± SE; 7 biological replicate pools.

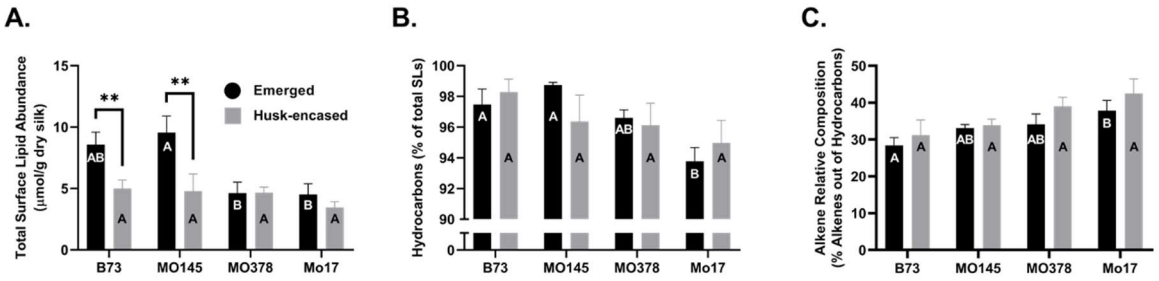


Figure 3 | Surface lipid compositions across four genotypes and two husk-encasement statuses. (**A**)Total surface lipids and relative abundances of (**b**) hydrocarbons (% hydrocarbons of total SLs), and (**c**) alkenes (% alkenes of total HCs). (**A-C**) Surface lipid traits for emerged and husk-encased silks were compared within each genotype (t-tests; *, $p<0.1$; **, $p<0.05$; and ***, $p<0.01$). Different uppercase letters above data-bars for total surface lipid abundance or specific lipid class denote statistically significant differences among genotypes from emerged or husk-encased silks (black and white letters respectively), at $p<0.05$, according to a Tukey's HSD test. Mean ± SE; 7 biological replicate pools.

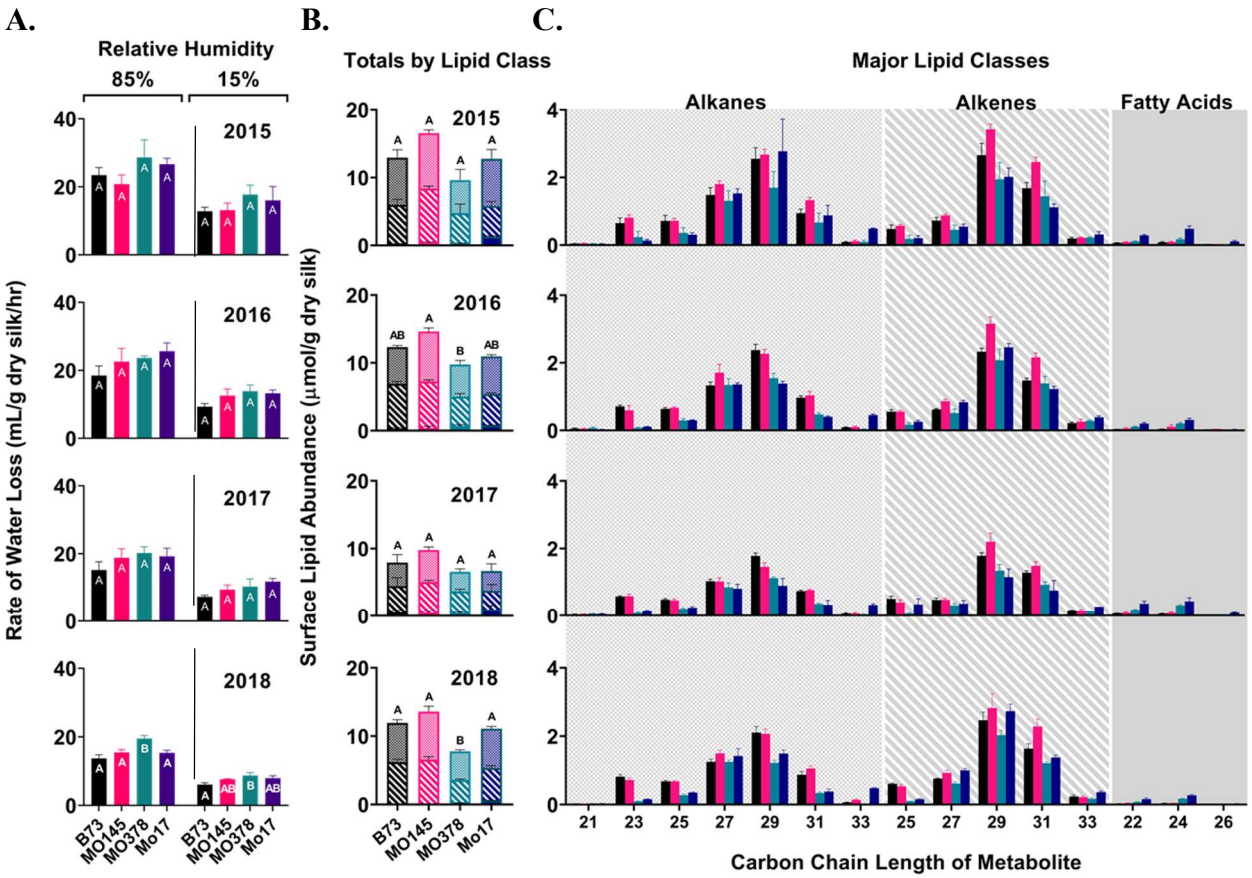


Figure 4 | Rates of water loss and surface lipid metabolome composition for emerged silks across the 4-year study. Emerged silk samples from four inbred genotypes, B73 (Black), MO145 (Pink), MO378 (Teal) and Mo17 (Navy blue), were assayed for (**a**) rates of water loss at two relative humidities, (**b**) total surface lipid accumulation, and (**c**) accumulation of individual constituents within specific surface lipid classes (i.e. alkanes, alkenes, and fatty acids). Alkanes are checkered, alkenes are striped, and fatty acids are solid. The most abundant metabolites from each class are plotted by chain-length. (**A-B**) Different uppercase letters within data-bars for rates of water loss or above data-bars for total surface lipid abundance denote statistically significant differences among genotypes within a specific year and at 15% or 85% relative humidity for rates of water loss, at $p<0.05$, according to a Tukey's HSD test. Mean ± SE; 4 biological replicate pools.

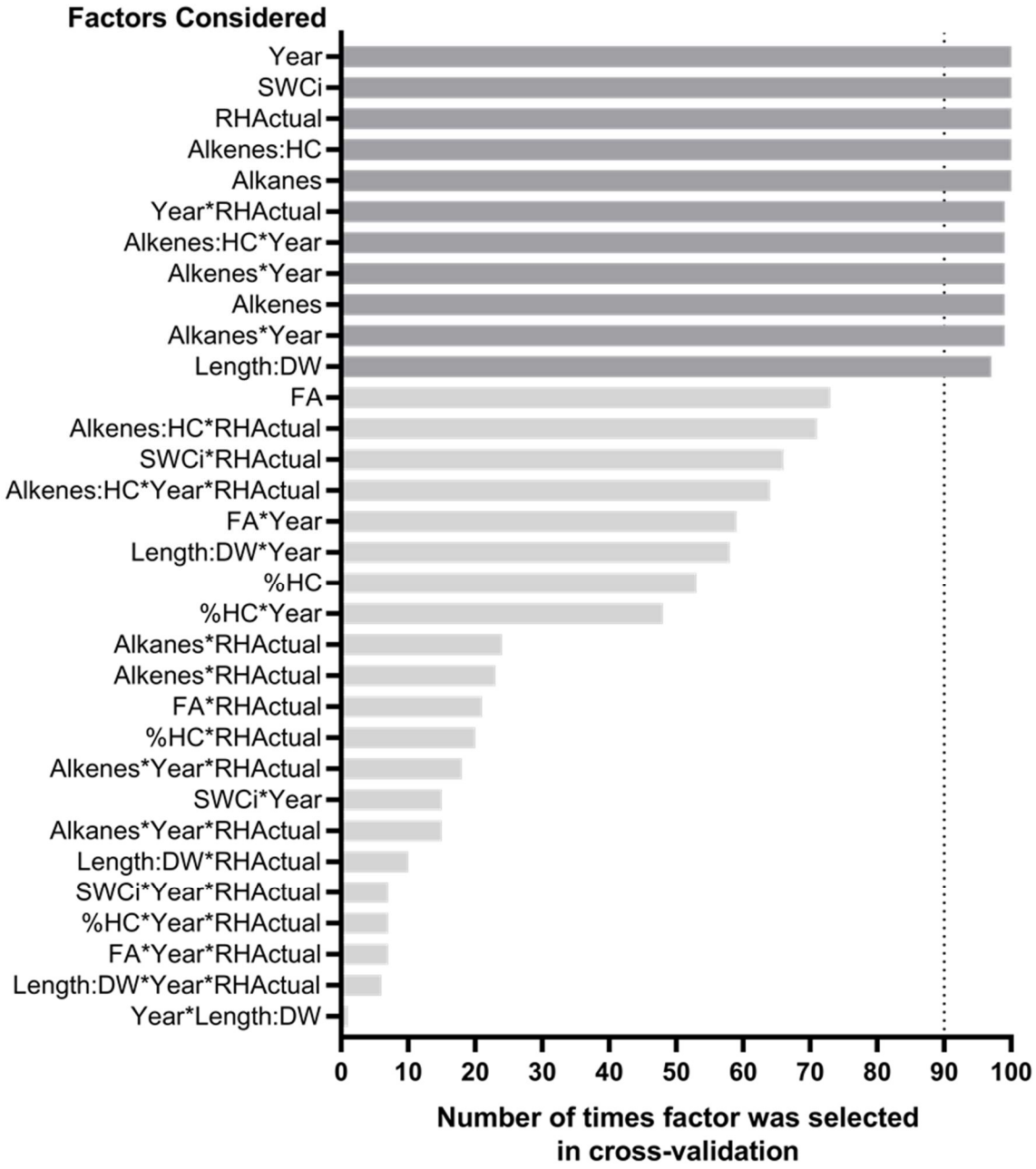


Figure 5 | Factor selection from ten rounds of 10-fold cross-validation. Factors listed on the Y-axis were considered for their impact on rates of water loss in the four year-dataset. The total number of times each factor was selected via 10 iterations of 10-fold cross-validation is represented by bars along the X-axis. Factors that were selected more than 90 times (threshold indicated by dotted line) are displayed in a darker shade of grey. SWCi, silk water content; relative humidity, relative humidity; HC, hydrocarbons; DW, dry weight; FA, fattty acids.

Tables

Table 1. Multi-factor ANOVA for rates of water loss from silk tissue.

Fixed Effects	**F-value**	**Partial R^2**	**p-value**[a]
Relative Humidity	$F_{1,223}$= 275.789	0.333	< 0.0001
Temperature	$F_{1,223}$= 200.421	0.242	< 0.0001
Genotype	$F_{3,223}$= 19.407	0.07	< 0.0001
Husk-Encasement Status	$F_{1,223}$= 46.091	0.056	< 0.0001
Genotype X Relative Humidity	$F_{3,223}$= 5.172	0.019	0.002
Genotype X Husk-Encasement Status	$F_{3,223}$= 3.283	0.012	0.022
Husk-Encasement Status X Temperature	$F_{1,223}$= 10.20	0.012	0.002
Temperature X Relative Humidity	$F_{1,223}$= 5.782	0.007	0.017
Genotype X Temperature	$F_{3,223}$= 1.405	0.005	ns
Genotype X Temperature X Relative Humidity	$F_{3,223}$= 0.705	0.003	ns
Genotype X Husk-Encasement Status X Temperature X Relative Humidity	$F_{3,223}$= 0.726	0.003	ns
Genotype X Husk-Encasement Status X Temperature	$F_{3,223}$= 0.432	0.002	ns
Husk-Encasement Status X Relative Humidity	$F_{1,223}$= 2.066	0.002	ns
Genotype X Husk-Encasement Status X Relative Humidity	$F_{3,223}$= 0.535	0.002	ns
Husk-Encasement Status X Temperature X Relative Humidity	$F_{1,223}$= 0.083	0	ns
Whole Model	**$F_{31,230}$ = 20.498**	**0.768**	**< 0.0001**

[a]Multi-factor ANOVA tested the main effects of genotype, husk-encasement status, temperature, relative humidity and the full-factorial set of interactions between these factors. Partial R^2 values indicate the proportion of variance explained by a given factor

Table 2. Rank of factors included in the optimized model according to the relative importance to rates of water loss, as suggested by Bayesian model testing.

Optimized model Rate of water loss ~ Relative Humidity + silk water content + growing year + %hydrocarbons + % alkenes + (% alkenes x growing year) + silk length ratio		
Factor	**Bayes factor**	**Sign of regression coefficients**
Relative humidity observed (RHActual)[a]	**2.17×10^{26}** [b]	**Negative** [c]
Alkanes (Alka)	**46.9**	**Negative**
Initial silk water content (SWC_i)	**38.7**	**Positive**
Year	**27.2**	**Negative**
Percent Alkenes over total hydrocarbons (Alkanes:HC)	2.15	Negative
Alke:HC x Year	2.15	Positive
Silk length per dry weight (Length:DW)	0.26	Positive

[a] The factors are ranked from high to low based on the Bayes factors.
[b] The factors with a Bayes factor >10 (numbers in bold) exhibit strong impact on rates of water loss according to Bayesian hypothesis testing analysis.
[c] Regression coefficients are estimated by fitting the final model to 4-year combined dataset and the signs of coefficients indicates the direction of influence. A positive coefficient represents a positive association between rates of water loss and the given factor and a negative coefficient represents a negative association with the given factor. For example, increased relative humidity is associated with a decrease in rate of water loss.

Table 3. Application of the optimized model to a new dataset.

Optimized model (excluding growing year)[a] Rate of water loss ~ Relative Humidity + silk water content + %hydrocarbons + % alkenes + silk length ratio			
Factor	**F-value**	**Partial R^2**	**p-value**
Relative humidity observed (RHActual)	$F_{1,71} = 117.96$	0.518	<0.0001
Initial silk water content ($SWC_{initial}$)	$F_{1,71} = 17.98$	0.079	<0.0001
Alkanes (Alka)	$F_{1,71} = 5.09$	0.022	0.0273
Percent Alkenes over total hydrocarbons (Alkanes:HC)	$F_{1,71} = 1.57$	0.007	ns
Silk length per dry weight (Length:DW)	$F_{1,71} = .0042$	0	ns
Whole Model	**$F_{5,71} = 32.35$**	**0.626**	<0.0001

[a]Growing year was not included as a factor in the model because the validating dataset was obtained from a single growing year (2016).

Supplement

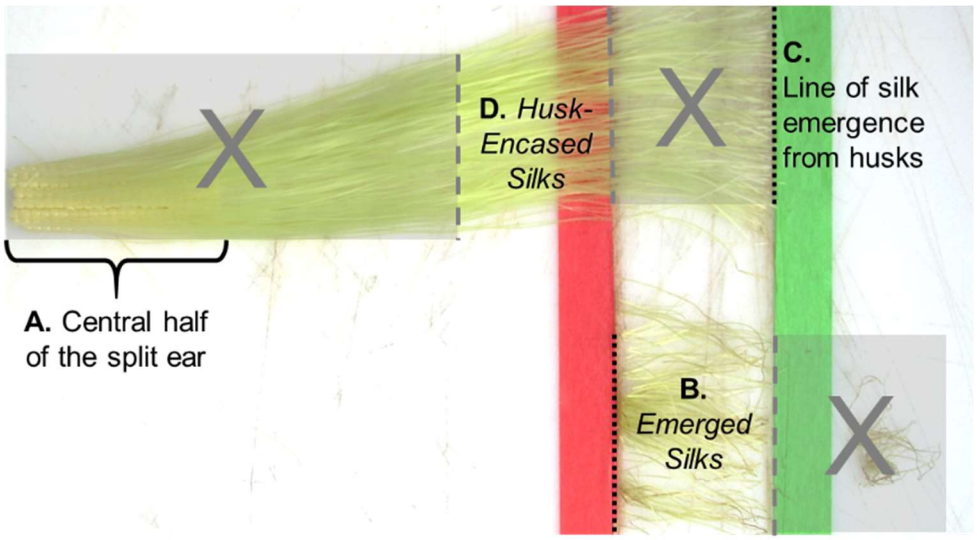


Figure S1 | Schematic of silk tissue processing. Harvested ears were cut along the length of the ear and separated into two long sections. The bottom and top quarters of the cob and associated silks were snapped off and discarded, leaving the central half of the ear to be laid out on a cutting board. Emerged silks were cut as 3 cm sections starting at the line of silk emergence and silks above this line were discarded (indicated by Xs). The 3 cm segment proximal to the emerged silks was also discarded and the subsequent 3 cm section was designated as husk-encased silks.

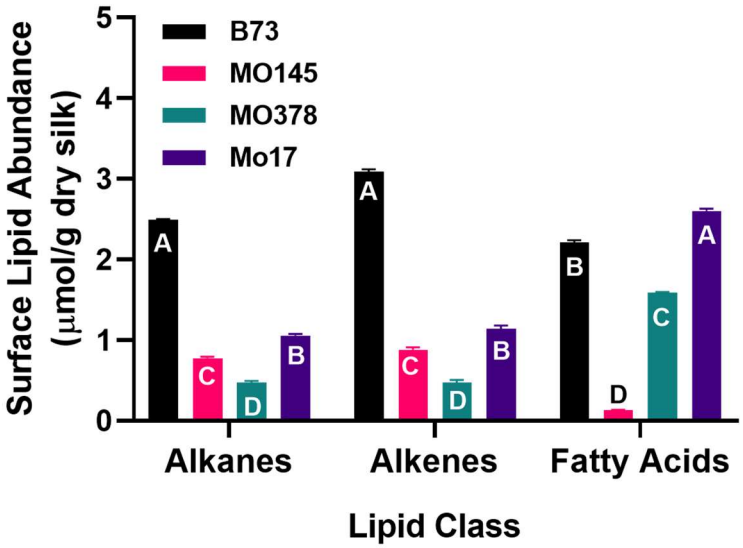


Figure S2 | Lipid profiles of the IBMRILs MO145 and MO378, and the parental inbred lines. Surface lipid classes were quantified from emerged silks, 3-days after silks first emerged from the encasing husk leaves. Different uppercase letters above data-bars for each specific lipid class denote statistically significant differences among genotypes, at $p<0.05$, according to a Tukey's HSD test. Mean ± SE; 7 biological replicate pools.

Table S1. ANOVA for rates of water loss calculated from silks collected in growing year 2015 to estimate the impact of physical silk traits (including surface lipid traits), temperature, and relative humidity.

Factor	**F-value**	**Partial R^2**	**p-value**
Relative humidity observed (RH)	$F_{1,223}$= 307.4	0.328	< 0.0001
Temperature (Temp)	$F_{1,223}$= 223.3	0.238	< 0.0001
Silk length per dry weight (Length:DW)	$F_{1,223}$= 70.54	0.075	< 0.0001
Initial silk water content ($SWC_{initial}$)	$F_{1,223}$= 49.49	0.053	< 0.0001
Percent hydrocarbons (%HC)	$F_{1,223}$= 15.69	0.017	< 0.0001
Temp x RHActual	$F_{1,223}$= 15.49	0.017	< 0.0001
Alkanes (Alka)	$F_{1,223}$= 13.50	0.014	< 0.001
Length:DW x RHActual	$F_{1,223}$= 9.180	0.01	0.003
Fatty acids (FA)	$F_{1,223}$= 8.604	0.009	0.004
$SWC_{initial}$ x Temp	$F_{1,223}$= 7.171	0.008	0.008
$SWC_{initial}$ x RHActual	$F_{1,223}$= 7.699	0.008	0.006
Length:DW x Temp	$F_{1,223}$= 4.939	0.005	0.027
Alka x RHActual	$F_{1,223}$= 2.426	0.003	ns
FA x RHActual	$F_{1,223}$= 3.174	0.003	ns
FAs x Temp	$F_{1,223}$= 1.511	0.002	ns
Alkene:HC ratios (Alke:HC)	$F_{1,223}$= 0.881	0.001	ns
Alkenes (Alke)	$F_{1,223}$= 0.480	0.001	ns
%HC x Temp	$F_{1,223}$= 0.562	0.001	ns
Alka x Temp	$F_{1,223}$= 0.880	0.001	ns
%HC x RHActual	$F_{1,223}$= 1.168	0.001	ns
Alke:HC x Temp	$F_{1,223}$= 0.001	0	ns
Alke x Temp	$F_{1,223}$= 0.008	0	ns
Alke:HC x RHActual	$F_{1,223}$= 0.097	0	ns
Alke x RHActual	$F_{1,223}$= 0.041	0	ns
%HC x Temp x RHActual	$F_{1,223}$= 0.297	0	ns
Alke:HC x Temp x RHActual	$F_{1,223}$= 0.228	0	ns
Alka x Temp x RHActual	$F_{1,223}$= 0.131	0	ns
Alke x Temp x RHActual	$F_{1,223}$= 0.253	0	ns
FA x Temp x RHActual	$F_{1,223}$= 0.104	0	ns
$SWC_{initial}$ x Temp x RHActual	$F_{1,223}$= 0.027	0	ns
Length:DW x Temp x RHActual	$F_{1,223}$= 0.012	0	ns

A. **B.**

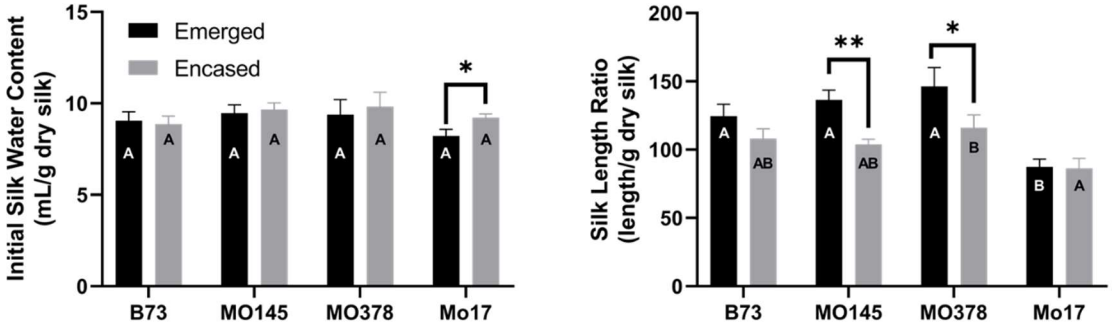


Figure S3 | Non-lipid silk traits across four genotypes and two husk-encasement statuses. (**A**) Silk water content and (**b**) silk length, normalized to dry silk mass. (**A-B**) Emerged and husk-encased silks were compared within each genotype using a t-test (*, $p<0.1$; **, $p<0.05$; and ***, $p<0.01$). Different uppercase letters above data-bars for a specific physical silk trait denote statistically significant differences among genotypes from emerged or husk-encased silks (black and white letters respectively), at $p<0.05$, according to a Tukey's HSD test. Mean ± SE; 7 biological replicate pools.

Table S2. ANOVA of rates of water loss from emerged silks collected across the 4-year study. Three-way ANOVA assessed the effects of genotype, relative humidity and growing year on rates of water loss.

Factor	**F-value**	**Partial R^2**	**p-value**
Relative Humidity (RH)	$F_{1,127} = 64.36$	0.267	**<0.0001**
Year	$F_{3,127} = 16.15$	0.201	**<0.0001**
Genotype	$F_{3,127} = 5.728$	0.071	**0.0012**
Year*Genotype	$F_{9,127} = 1.41$	0.053	0.1944
Year*Genotype*RH	$F_{9,127} = 0.171$	0.006	0.9965
Genotype*RH	$F_{3,127} = 0.196$	0.002	0.8989
Year*RH	$F_{3,127} = 0.165$	0.002	0.9196

Table S3. ANOVA of total surface lipid accumulation from emerged silks collected across the 4-year study. Two-factor ANOVA assessed the effects of genotype and year on total surface lipid accumulation.

Factor	**F-value**	**Partial R^2**	**p-value**
Year	$F_{3,63} = 9.00$	0.254	**<0.0001**
Genotype	$F_{3,127} = 5.73$	0.257	**<0.0001**
Year*Genotype	$F_{9,127} = 1.41$	0.037	0.9054

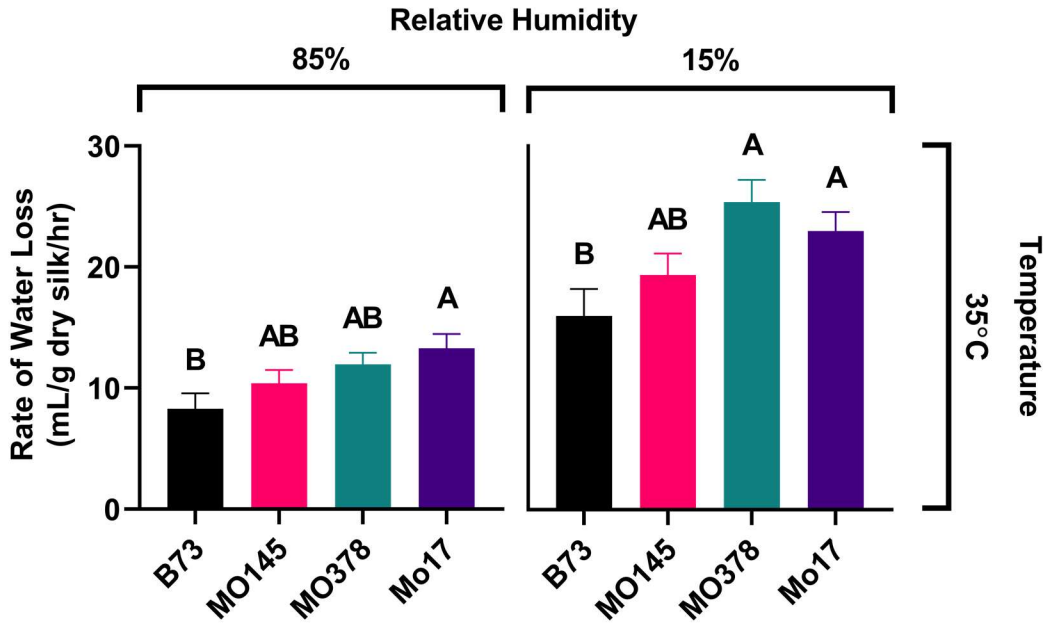


Figure S4 | Rates of water loss from silks exposed to different relative humidities. Emerged silks from four genotypes harvested in growing year 2016 were split amongst two relative humidities at 35˚C and measured for rates of water loss. Different uppercase letters above data-bars from each relative humidity denote statistically significant differences among genotypes, at $p<0.05$, according to a Tukey's HSD test. Means ± SE; 9 biological replicate pools.

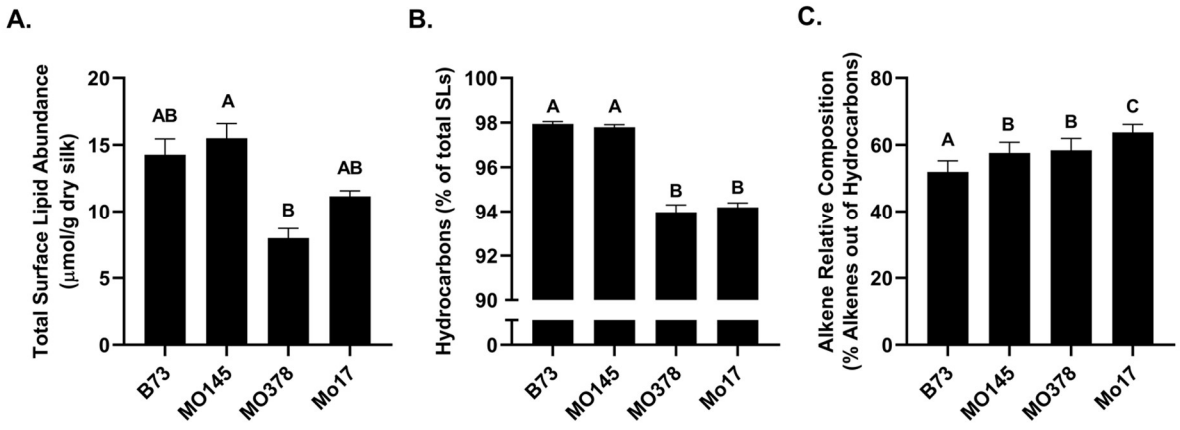


Figure S5 | Cuticular lipid profiles for emerged silks sampled in 2016. Emerged silks corresponding to the samples that were split amongst treatments and plotted in **Figure S4** were extracted, derivatized and profiled via GC-FID. (**A**) Total surface lipid accumulation is shown. Relative abundances of (**b**) hydrocarbons relative to total SLs, and (**c**) alkenes relative to total hydrocarbons, were calculated from individual metabolites and represented as percentages. Different uppercase letters above data-bars denote statistically significant differences among genotypes, at $p<0.05$, according to a Tukey's HSD test. Means ± SE; 9 biological replicate pools.

Table S4. ANOVA for rates of water loss to estimate the impact of factors selected by 10 iterations of 10-fold cross-validation

Factor	**F-value**	**Partial R^2**	**p-value**
Relative humidity observed (RHActual)	$F_{1,127}$= 254.4	0.496	< 0.0001
Year	$F_{1,127}$= 82.28	0.16	< 0.0001
Silk length per dry weight (Length:DW)	$F_{1,127}$= 18.42	0.036	< 0.0001
Initial silk water content ($SWC_{initial}$)	$F_{1,127}$= 17.42	0.034	< 0.0001
Alkenes:HC ratios (Alke:HC)	$F_{1,127}$= 7.004	0.014	0.00926
Alke:HC x Year	$F_{1,127}$= 5.323	0.01	0.0228
Alkanes	$F_{1,127}$= 4.794	0.009	0.0306
Alkenes x Year	$F_{1,127}$= 3.845	0.007	ns
Year x RHActual	$F_{1,127}$= 3.599	0.007	ns
Alkenes	$F_{1,127}$= 0.0172	0	ns
Alkanes x Year	$F_{1,127}$= 0.087	0	ns

Table S5. Non-lipid physical silk trait measurements of emerged silks, including initial silk water content and silk length relative to dry weight. Different uppercase letters within parentheses denote statistically significant differences amongst genotypes for a given trait on a given year, at $p<0.05$, according to a Tukey's HSD test. Means ± SD; four biological replicates pools.

Non-lipid physical silk trait	**Year**	**B73**	**MO145**	**MO378**	**Mo17**
Initial silk water content (mL/g dry silk) [a]	2015	10.2 ± 2 (A)	10.5 ± 2.1 (A)	9.5 ± 3.3 (A)	8.5 ± 1.8 (A)
	2016	8.1 ± 0.8 (A)	8.3 ± 1 (A)	7.8 ± 0.5 (A)	7.7 ± 0.8 (A)
	2017	6.8 ± 0.9 (A)	9 ± 0.9 (B)	7.9 ± 1.2 (AB)	7.9 ± 0.6 (AB)
	2018	8 ± 0.9 (A)	10 ± 0.4 (B)	9 ± 0.8 (AB)	8.6 ± 0.7 (AB)
Silk length relative to dry weight (m/g dry silk) [a]	2015	137 ± 31 (A)	160 ± 30 (A)	158 ± 59 (A)	98 ± 12 (A)
	2016	110 ± 21 (A)	116 ± 16 (A)	106 ± 3 (AB)	77 ± 9 (B)
	2017	114 ± 25 (A)	115 ± 10 (A)	122 ± 15 (A)	88 ± 13 (A)
	2018	104 ± 15 (A)	122 ± 16 (AB)	97 ± 9 (B)	81 ± 5 (B)

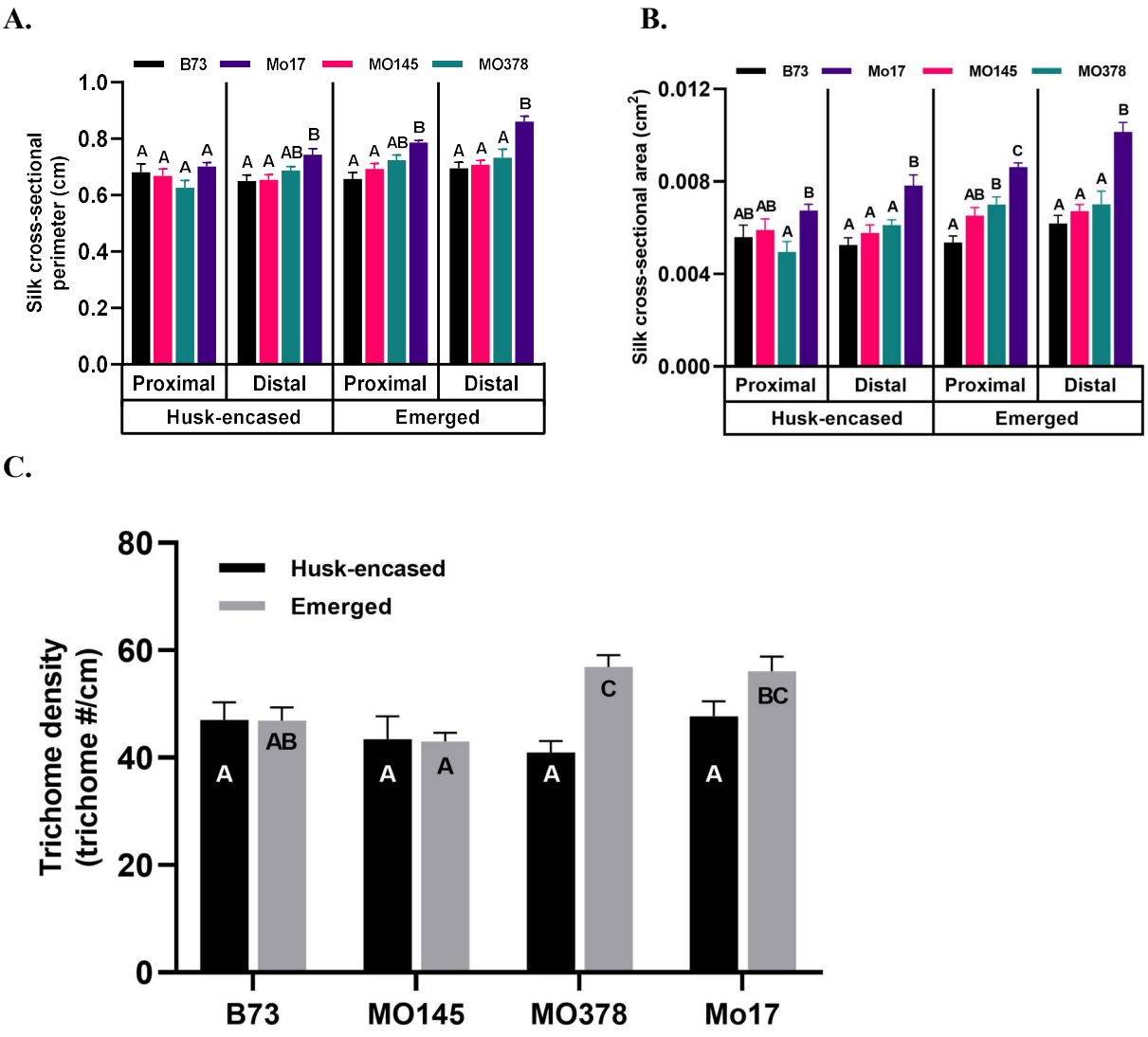


Figure S6 | Silk morphological features. Silks from four genotypes harvested in growing year 2017 were cut and imaged on both proximal and distal sides of emerged and husk-encased silk segments to obtain cross-sectional (**a**) perimeters and (**b**) areas at each position. **C**) Trichomes were also counted across silk segments to determine trichome density per cm of silk length for both emerged and husk-encased segments. Different uppercase letters within data-bars denote statistically significant differences among genotypes at a given position along the silk length, at $p<0.05$, according to a Tukey's HSD test. Means ± SE; 5-10 biological ear replicates.

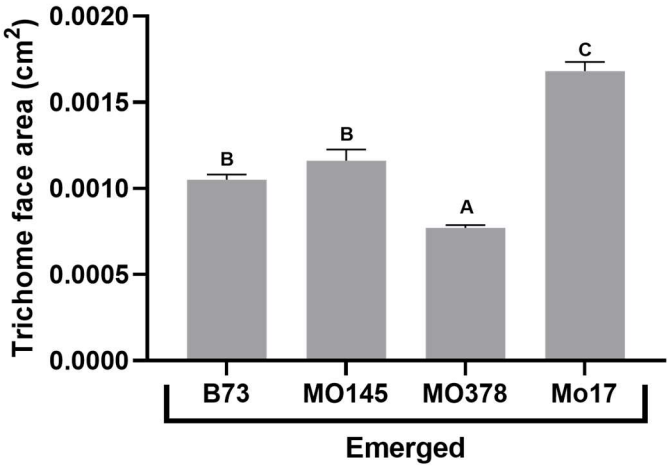


Figure S7 | Emerged silk trichome size by genotype. Silks from four genotypes harvested in growing year 2016 were cut and imaged to obtain trichome face areas, the 2-dimensional side view image of a trichome. Different uppercase letters above data-bars denote statistically significant differences amongst genotypes, at $p<0.05$, according to a Tukey's HSD test. Means ± SE; 10-15 biological ear replicates.

CHAPTER 4. COMPLEMENTATION OF THE ARABIDOPSIS CER3 MUTANT WITH GL1-LIKE2 AND IDENTIFICATION OF A CRITICAL AMINO ACID REGION

Bri Vidrine[1], Korinna Radke[2], Libuse Brachova[1], Liza Alexander[1], Jennifer Chmielowski[1], Marna Yandeau-Nelson[2], & Basil Nikolau[1]

[1]Iowa State University, Department of Biochemistry Biophysics & Molecular Biology, Ames, IA, and USA

[2]Iowa State University, Department of Genetics, Development & Cell Biology, Ames, IA, and USA

Key words: Surface lipids, Cuticular wax, Maize, *Arabidopsis*, Glossy, CER3

Abstract

Aliphatic hydrocarbons, including alkanes and alkenes, are an important class of cuticular waxes that protect plants from water stress and eventually become a core component of fossil fuels. Six maize genes share homology with both the *CER1* and *CER3* in Arabidopsis, which are proposed to function as a fatty acyl decarbonylase and fatty acyl reductase respectively, catalyzing the final steps of alkane biosynthesis. Five of these maize homologs functionally complement either the *cer1* or *cer3* mutants. A synthetic version of the *ZmGl1-like2* maize homolog *ZmGl1-like2*_T01 (i.e. GRMZM2G029912_T01), previously did not functionally complement the *cer1* and *cer3* knockout lines. In this study, heterologous overexpression of two native *ZmGl1-like2* transcripts isolated from maize inbred lines B73 and Mo17 functionally complemented the Arabidopsis *cer3-8* mutant. The synthetic GRMZM2G029912_T01 allele was found to include a 13-amino acid insertion between amino acids 91-93 of the protein sequence, N-terminal to the putative fatty acid hydroxylase domain. Therefore, this region of the functional protein is critical for the *zmGL1-like2* protein to function in cuticular lipid synthesis.

Introduction

Cuticular surface lipids coat aerial plant tissues and are the first line of defense against both biotic and abiotic stresses (Shepherd & Wynne Griffiths, 2006). One class of these surface lipids are aliphatic hydrocarbons, including alkanes and alkenes, which are hydrophobic, high-energy compounds and core components of fossil fuels. Despite their industrial value, characterization of alkane biosynthesis has proven challenging; many proteins involved in this pathway are membrane-bound and thus difficult to characterize within heterologous systems (Yeats & Rose, 2013). Phenotypic studies have identified mutants that are deficient in their cuticular surface lipid coating in maize (*glossy* mutants) and *Arabidopsis* (*eceriferum* mutants), among others including rice (Post-Beittenmiller, 1998; Schnable et al., 1994; X. Zhou et al., 2015). Metabolite analysis of lipids extracted from the surfaces of these mutants has identified a few mutations which cause dramatic reductions in alkane accumulation. The *Arabidopsis CER1* and *CER3* genes were of interest because the stems of the corresponding mutants accumulate higher than normal levels of primary alcohols. In the *cer1* mutant, higher levels of aldehydes are observed with a concomitant reduction in alkanes and predicted downstream lipid products (*i.e.* ketones and secondary alcohols) from the alkane producing pathway. Additionally, *CER1* overexpression was shown to increase very-long chain alkanes above wildtype levels (Bourdenx et al., 2011). In 2012, Bernard et al. demonstrated that these genes were directly involved in the terminal steps of the alkane biosynthetic pathway using a *Saccharomyces cerevisiae* host (Bernard et al., 2012), leading to the proposal that CER1 and CER3 form a complex that reduces very-long-chain fatty acid precursors (>29 carbon atoms in length) into corresponding aldehyde intermediates that are subsequently decarbonylated to the alkane by CER1. The hemoprotein, CYTB5, which is thought to donate electrons to CER1 enhanced nonacosane production, however, it was not essential for alkane production. More recently, an *Arabidopsis* CER1

homolog, CER1-like1, was also shown to similarly function alongside CER3 and CYTB5 within yeast (Pascal et al., 2019). Collectively, these results suggest that both *CER1* and *CER3* are involved in the final steps of alkane biosynthesis (Matthew A. Jenks et al., 1995).

Although progress has been made toward understanding plant alkane biosynthesis, the genes responsible for the final steps of this pathway in monocots have not been fully explored. Approximately 30 *glossy* genes have been phenotypically identified in seedlings, at which stage cuticular lipids from wildtype plants comprise aldehydes and alcohols and are relatively deplete of alkanes (Avato, Bianchi, & Salamini, 1985; G. Bianchi, Avato, & Salamini, 1979; Li et al., 2018; Moose & Sisco, 1994). The maize *glossy1* mutant, identified by its shiny or 'glossy' phenotype and its ability to accumulate beads of water relative to the wildtype (Hayes & Brewbaker, 1928), was shown to have reduced lipid accumulation on seedling leaves (A. Bianchi, Bianchi, Avato, & Salamini, 1985). Although the seedling chemotype did not directly associate *GLOSSY1* with alkane biosynthesis, homology to the CER1 and CER3 proteins suggests that *GLOSSY1* and/or *GLOSSY1* paralogs are responsible for the terminal steps of alkane biosynthesis. Partial complementation was observed when expressing maize *Glossy1* in an *Arabidopsis cer3* mutant but not in the *cer1* mutant (Chmielowski, 2016), suggesting that maize GLOSSY1 functions like CER3. In the same study, five *Glossy1* homologs were identified within the B73 maize genome, named *zmGl1-like1* through *zmGl1-like5*. Three of these genes (*ZmGl1-like3*, *zmGl1-like4*, and *ZmGl1-like5*) were shown to either fully, or in the case of *ZmGl1-like3,* partially complement the *cer1* mutant. In contrast, *ZmGl1-like1* complemented the *cer3* mutant. Although *ZmGl1-like2* did not complement either *Arabidopsis* gene, overexpression exacerbated the primary alcohol chemotype of the *cer3* mutant (Chmielowski, 2016).

Because the function of *zmGL1-like2* in the hydrocarbon biosynthetic pathway remains unresolved, this study sought to further investigate its function through heterologous over-expression of three *ZmGl1-like2* alleles (*ZmGl1-like2_T01* derived from version 3 of the maize B73 draft genome sequence, and cloned alleles from maize inbred lines, B73 and Mo17) into the *Arabidopsis cer3* and *cer4* mutants, and the wildtype ecotype, Col-0. Over-expression of *zmGL1-like2_T01* in the *cer4* mutant did not affect production of primary alcohols or impact the composition of other surface lipids. Two *ZmGl1-like2* alleles isolated from cDNA from inbred lines B73 and Mo17 complemented the *cer3* mutant and failed to complement the *cer1* mutant. *zmGL1-like2_T01,* which harbors a 13-amino acid insertion, failed to complement either the *cer1* or *cer3* mutants. Collectively, this study reveals that *zmGL1-like2* is a functional homolog of *Arabidopsis* CER3 and is a likely a key component of a putative alkane-forming complex in maize.

Results

zmGL1-like2 does not functionally complement the CER4 alcohol reductase

Primary alcohol accumulation was observed in response to *zmGL1-like2_T01* overexpression in both WT and *cer3-8* lines (Chmielowski, 2016), suggesting that it could be responsible for primary alcohol formation. To test the hypothesis that *zmGL1-like2* is a functional homolog of the *Arabidopsis* alcohol-forming fatty acid reductase CER4 (Rowland et al., 2006), a complementation experiment was conducted using a synthetic *ZmGl1-like2_T01* construct derived from version 3 of the draft genome sequence for maize inbred line, B73. *ZmGl1-like2_T01* under the control of the 35S promoter was introduced into two distinct *cer4Δ/cer4Δ* Arabidopsis lines, *cer4-3*, and *cer4-575*. Plants were grown to maturity and surface lipid composition was profiled from stems, revealing that the *ZmGl1-like2_T01* allele did not

increase chemotypic expression of primary alcohols in either of the *cer4* mutant lines (**Fig. 1 & S1**). The only detectable differences in surface lipid composition from plants overexpressing *ZmGl1-like2_T01* were in the *cer4-3* mutant, a 0.2% and 2.1% reduction in the accumulation of the C30 aldehyde and the C28 alkane, respectively (**Fig. 1**). Therefore, it is unlikely that the *ZmGl1-like2_T01* impacts primary alcohol synthesis or accumulation. However, *ZmGl1-like2* overexpression was not confirmed, therefore it is possible that *ZmGl1-like2* was transcriptionally silent despite its confirmed presence within the genome.

Amino acid sequence differences exist amongst two genuine *ZmGl1-like2* alleles and the synthetic *ZmGl1-like2-T01* allele

Alignment of the predicted protein sequence of *zmGL1-like2*_T01 derived from the draft genome sequence (B73_V3 MaizeGDB) to *Arabidopsis CER3* revealed 13 extra amino acids inserted between amino acids 91-93 of the *zmGL1-like2_T01* allele relative to CER1 and CER3 and the other five maize homologs (**Fig. 2A**). To assess whether a different isoform of *zmGL1-like2* is present in the maize transcriptome, *ZmGl1-like2* transcripts were sequenced from silks of inbred lines, B73 and Mo17 and compared to the synthetic *ZmGl1-like2_T01* allele from version 3 of the draft genome sequence. B73 and Mo17 were specifically selected because B73 has been shown to accumulate higher levels of alkanes within maize silks relative to the Mo17 inbred line (Dennison et al., 2019; Loneman et al., 2017; Perera et al., 2010). In case protein differences accounted for altered alkane production, both B73 and Mo17 alleles were cloned from silk cDNA. Sequencing and subsequent alignment with *ZmGl1-like2_T01* and the *Arabidopsis* and maize homologs revealed amino acid differences between the B73 and Mo17 transcripts within the predicted fatty acid hydroxylase domain of *zmGL1-like2*, suggesting that there could be altered protein function between the two alleles (**Fig. 2A**). In addition, the genuine transcripts from B73 and Mo17 did not contain the additional 13-amino acids contained in *zmGL1-*

like2_T01 (**Fig. 2A**). Therefore, it was hypothesized that the *zmGL1-like2_B73* and *zmGL1-like2_Mo17* would be functional, unlike *zmGL1-like2_T01*.

The B73 transcript isolated from silk cDNA matched the primary transcript from the new version of the draft genome (Zm00001d046865_T001 from Maize B73 RefGen_v4) and was thus used to conduct a phylogenetic analysis relative to *Arabidopsis* and maize homologs. Two clades were identified: clade I contains *Arabidopsis* CER3, *GLOSSY1, zmGL1-like1* and *zmGL1-like2*, and clade II contains *Arabidopsis* CER1 and *zmGL1-like3*, *zmGL1-like4* and *zmGL1-like5* (**Fig. 2A**). *zmGL1-like2* and CER3 share 75% amino acid similarity, suggesting that *zmGL1-like2* is more likely to share function to CER3 as opposed to CER1 (52% amino acid similarity; **Table S1**). Moreover, *zmGL1-like2* and CER3 show high homology within the putative conserved WAX2_C domain, which remains uncharacterized. Notably, CER1 shows little sequence homology with either *zmGL1-like2* or CER3 within this region (**Fig. 2B**).

The *zmGL1-like2* alleles from inbreds B73 and Mo17 functionally complement the *Arabidopsis cer3* knockout mutant

To test for functional complementation of the *cer3* mutant in *Arabidopsis*, all three *zmGl1-like2* alleles (B73, Mo17, and T01) were heterologously overexpressed under the 35S promoter in both the *Arabidopsis* wildtype Col-0 ecotype and the Arabidopsis *cer3-8/cer3-8* mutant lines. After confirmation of *ZmGl1-like2* expression, surface lipid compositions from primary stems were assessed for *cer3* mutants and wildtype lines harboring one of the three *ZmGl1-like2* transgenes (B73, Mo17, or T01) and the untransformed controls (**Fig. 3**). Three independent transgenic events were evaluated for each *ZmGl1-like2* allele. Although some variation existed among the three unique transformation events per *ZmGl1-like2* allele, the total surface lipid concentration was elevated by approximately 2-fold relative to the *cer3-8/cer3-8* control for *ZmGl1-like2-Mo17* and *ZmGl1-like2-B73*, whereas the original *ZmGl1-like2_T01*

allele did not induce a change in surface lipid accumulation compared to the *cer3* mutant in any event. In a comparison with wildtype Col-0, both *ZmGl1-like2_B73* and *ZmGl1-like2_Mo17* restored levels of total aldehydes and hydrocarbons to wildtype levels (**Fig. 3**). Examination of individual surface lipid constituents reveals similar increases in accumulation of aldehydes and alcohols across carbon-chain lengths (i.e. C26, C28, C30) (**Fig. 4)**. The three aldehyde constituents increased by ~10-fold in lines expressing *ZmGl1-like2_B73* and *ZmGl1-like2_Mo17* relative to both the control and the *ZmGl1-like2_T01* allele. Increases in primary alcohols were more modest. In contrast, overexpression of both *ZmGl1-like2_B73* and *ZmGl1-like2_Mo17* induced accumulation of the C29 alkane with slightly more modest increases in the corresponding C29 ketone, which are major constituents of stem surface lipids from wildtype *Arabidopsis*. Collectively, these results demonstrate that the *ZmGl1-like2* alleles from both inbred lines functionally complement *cer3-8/cer3-8* surface lipid composition and total accumulation. Overexpression of the *ZmGl1-like2_T01* allele in the *cer3* mutant, however, had no effect on surface lipid composition as compared to the untransformed *cer3* mutant (**Fig. 4**), demonstrating that it does not functionally complement. Therefore, we conclude that the 13-amino acid sequence present in the T01 allele prevents Z*mGl1-like2* from functional complementation of *cer3*.

Overexpression of *ZmGl1-like2* in wildtype *Arabidopsis*

Each of the three *ZmGl1-like2* alleles were introduced into wildtype *Arabidopsis* to determine whether overexpression of *ZmGl1-like2* could alter the wildtype surface lipid phenotype. Three independent transgenic events were evaluated for each *ZmGl1-like2* allele (**Fig. 3**). Overexpression of each of the *ZmGl1-like2* alleles did not altered total surface lipid accumulation in any of the three transgenic events considered for the B73, Mo17, and T01 alleles as compared to untransformed wildtype Col-0. Repetition of this experiment again revealed no

differences in total surface lipids between the overexpression alleles and the control within the wildtype background. In addition, all individual lipid constituents accumulated to similar levels between transgenic strains and the untransformed wildtype control (**Fig. 5**). Therefore, overexpression of the *ZmGl1-like2* alleles did not change surface lipid expression in the wildtype *Arabidopsis* background.

***zmGL1-like2* does not complement the *Arabidopsis cer1* knockout background**

To test for functional complementation of the *Arabidopsis* gene *CER1*, *ZmGl1-like2_Mo17* and *ZmGl1-like2_B73* were introduced into the *Arabidopsis cer1-3* knock-out background driven under the 35S promoter. Expression of *ZmGl1-like2_B73* in the *cer1-3* knockout did not induce changes in surface lipid composition as compared to the *cer1-3* knockout alone. In contrast, expression of *ZmGl1-like2_Mo17* in the *cer1-3* knockout caused a complete absence of C28 and C30 aldehydes and a corresponding 2.8-fold increase in the C30 primary alcohol (**Fig. 6**). This result, however, was a relatively small effect and was observed in only one of the four transgenic events. Therefore, none of the transgenic lines showed complementation with *cer1-3*.

Discussion

Complementation tests with three different alleles of *zmGL1-like2* revealed that *zmGL1-like2* is a functional homolog to CER3. Overexpression of the B73 and Mo17 maize alleles of *ZmGl1-like2* within the *cer3-8* knockout *Arabidopsis* line demonstrated that both B73 and Mo17 alleles can functionally complement the *cer3-8* chemotype. These findings suggest that *ZmGl1-like2* functions to reduce very long chain acyl-CoAs in maize, as has been proposed for *CER3* function (Bernard et al., 2012; Matthew A. Jenks et al., 1995). Interestingly, the synthesized allele, *Zm-Gl1-like2_T01* did not function in this capacity. Because the only difference between *zm-GL1-like2_T01* and *zm-GL1-like2_B73* proteins were 13 extra amino acids inserted between

residues 91-92 relative to the genuine B73 allele, these 13 amino acids prevented *zmGL1-like2* from functioning like CER3. Because these sequences fall outside the WAX2_c and fatty acid hydroxylase domains, these added amino acids may affect protein tertiary structure by impacting folding or, because this gene has multiple transmembrane domains, it may interfere with the interaction between *zmGL1-like2* and the membrane. There is little support for the existence of *Zm-Gl1-like2_T01* in nature as the predicted transcript has since been removed from the newest version of the B73 reference genome (V4) due to insufficient RNA evidence (Jiao et al., 2017; Portwood et al., 2019). B73 and Mo17 *ZmGl1-like2* alleles function similarly across surface lipid classes and carbon-chain lengths, suggesting that the amino acid changes within the fatty acid hydroxylase domain are conservative.

The lack of complementation by *zmGL1-like2* in the *cer1* and *cer4* mutant backgrounds is consistent with the finding that *zmGL1-like2* functions as a CER3 functional homolog. Previous analyses of *cer1* mutants have demonstrated drastically reduced levels of alkanes and proposed downstream products of alkanes, including ketones and secondary alcohols, as compared to wildtype strains with functional copies of CER3 (Aarts et al., 1995; Hannoufa et al., 1993). All three overexpression constructs of *ZmGl1-like2* (B73, Mo17 & T01) did not complement this *cer1* mutant chemotype, because they did not cause changes to the levels of alkane, ketone or primary alcohol constituents and thereby could not restore the mutant phenotype of *cer1*. This negative result within the *cer1* background is consistent with the other previously tested, five maize homologs, which complemented either *cer1* or *cer3,* but not both (Chmielowski, 2016).

This study demonstrates that genuine versions of the B73 and Mo17 *ZmGl1-like2* alleles can functionally complement the *Arabidopsis cer3* mutant. Because of these results and the identification of other CER1 and CER3 functional homologs in maize and rice, there is a good

deal of support for the prediction that the maize hydrocarbon biosynthesis pathway acts in a similar manner to the pathway of *Arabidopsis* (Chmielowski, 2016; X. Zhou et al., 2015). The interaction between the redundant maize homologs remains to be explored. Replacing CER1, CER3 and CER1-like within the *Arabidopsis* system with different combinations of these mutants could help illuminate their potentially nuanced roles within the hydrocarbon biosynthesis pathway.

Methods

Plasmid Construction

Maize silk RNA was extracted from two inbred lines, B73 and Mo17, using TRIZOL (Rio, Ares, Hannon, & Nilsen, 2010). After DNase I (Invitrogen) treatment, transcripts were reverse transcribed into cDNA using SuperScriptIII and amplified for BP cloning into the pENTR/D-TOPO vector (Gateway), following manufacturer instructions. Primers were designed to amplify *zmGL1-like2* (GRMZM2G029912_T001) according to the B73 maize reference genome_V3 (Portwood et al., 2019; Schaeffer et al., 2011). Alleles were subsequently cloned from the entry vector into the *Arabidopsis* expression vector pEARLYGate100, which hosts a phosphinothricin resistance cassette for selection, using the Gateway system and/or In-Fusion cloning, per manufacturer instructions. Plasmids were confirmed via sequencing at the Iowa State University DNA Facility. A similar plasmid was constructed for *ZmGl1-like2_T01,* based on the B73 RefGen_v3 assembly of the maize genome (Gramene ID: GRMZM2G029912_T01; https://maizegdb.org/gene_center/gene/GRMZM2G029912). This sequence was codon optimized for expression in *Arabidopsis* using OptimumGene, and then synthesized and cloned into pMA-RQ vector by GenScript.

Arabidopsis lines and growth conditions

Arabidopsis was germinated in soil after three days of vernalization at 4°C. A controlled environment growth chamber was kept at a temperature of ~23°C with a relative humidity of ~35% and an average light intensity of 2500 lux. An *Agrobacterium tumefaciens* host (strain C58C1) was used to transfect *Arabidopsis* lines with pEARLYGate100 vectors carrying *zmGL1-like2* transgene alleles via the floral dip method (Clough & Bent, 1998). These *Arabidopsis* lines included wildtype Col-0 and T-DNA knockout lines in the Col-0 genetic background: *CER1/cer1-2* (SALK_014839), *CER3/cer3-8* (SALK_034318), *cer4-3/cer4-3* (SALK_038693C) or *cer4-575/cer4-575* (SALK_000575C) from the Arabidopsis Biological Resource Center (Berardini et al., 2015).

Seeds from transformed plants were harvested and screened via herbicide selection with 1:1000 phosphinothricin (Thompson et al., 1987). Wildtype and knockout *cer* alleles, and the heterologously expressed *zmGL1-like2* alleles were tested via PCR using primers detailed in **Table S1** to confirm genotypes of lines used in the analysis. *zmGL1-like2* overexpression was confirmed in the Col-0 wildtype and *cer3-8/cer3-8 Arabidopsis* lines by PCR amplification of cDNA templates derived from DNase-treated RNA.

Transgenic plants were transplanted after herbicide selection approximately one month after planting.

Surface lipid extraction and derivatization

Plants were harvested approximately 60 days after planting, when primary stems were 19 ± 2cm in length. The entire primary stem including leaves and flowers was removed and dipped for 1 minute in chloroform supplemented with the internal standards, methyl nonadecanoate (1.5 ug) and hexacosane (3 ug). Samples were dried under a stream of N^2. Extracts were silylated

with N,OBis(trimethylsilyl)trifluoroacetamide (Sigma-Aldrich) at 70°C for 30 min and then dried to completion under a stream of N^2. *Arabidopsis* lipid extracts were analyzed via GC-MS or GC-FID using the same gas chromatograph (Agilent Technologies Model 7890A) equipped with a DB-1 column. The chromatography program was as follows: the initial oven temperature was set to 120°C, ramped to 260°C by 10°C/min, held for 10 min, ramped to 320°C by 5°C/min, and held for 4 min to conclude the run. Peaks measured with MS were deconvoluted, identified and integrated using the AMDIS software package (Stein, 1999) in conjunction with in-house surface lipid compound libraries that were cross-checked with the NIST Mass spectral library (Acree & Chickos, 2018) and standards including a cocktail of fatty acid methyl esters (ME10-1KT; Sigma-Aldrich) and alkanes (49452-U; Supelco). Representative mutant and wildtype samples were evaluated on both GC-MS and GC-FID to produce a retention time-based conversion key that was used to assign FID peaks to specific metabolites. Peak areas were normalized using the hexacosane internal standard to determine the molar concentration of each constituent per tissue dry weight.

Phylogenetic and statistical analyses

A phylogenetic tree was generated via the maximum likelihood method using default conditions, using the software package, MEGA X (Kumar, Stecher, Li, Knyaz, & Tamura, 2018). Protein domain identification was based on NCBI DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST). All statistical analyses were carried out using JMP Pro 14. Dunnett's tests were used to compare treatments to a control and t-tests were used to make pair-wise comparisons. Samples were determined to be different from one another if the comparisons had a p-value less than a 0.05 threshold.

Acknowledgements

This work was facilitated by the W. M. Keck Metabolomics Research Laboratory at Iowa State University.

Figures

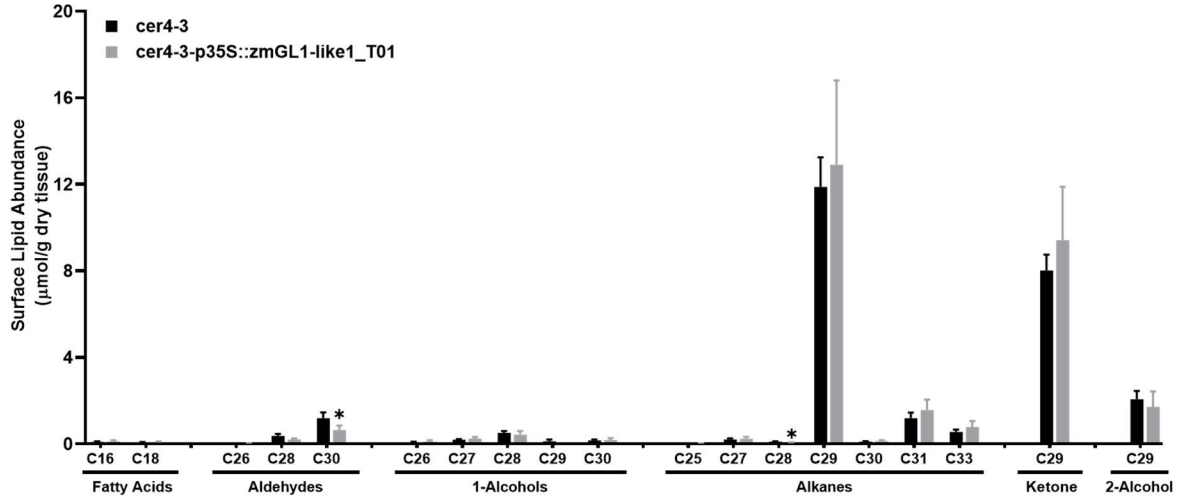
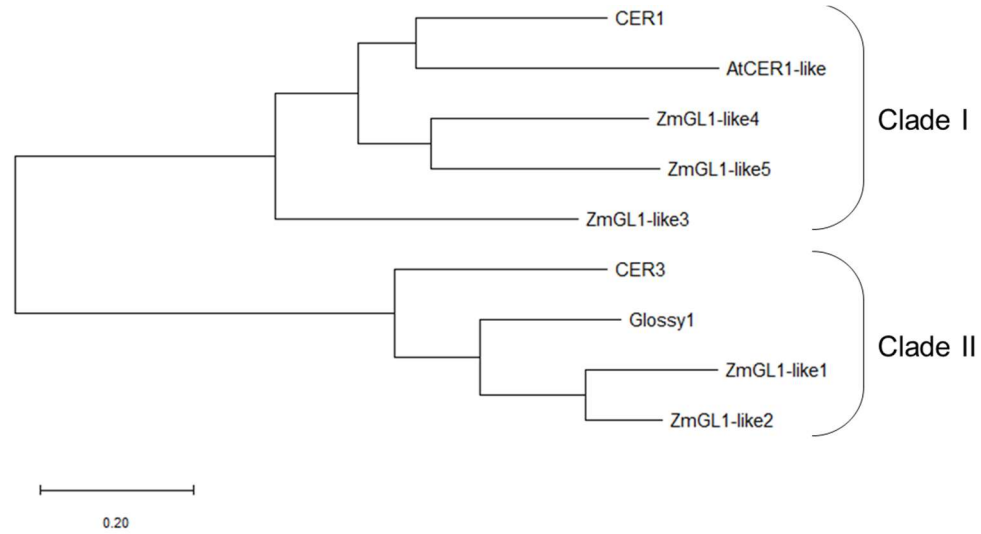


Figure 1 | Surface lipid composition on stems of Arabidopsis *cer4-3* mutant overexpressing *zmGl1-like1_T01*. *zmGl1-like2_T01* was overexpressed within the two *cer4-3* mutant *Arabidopsis* lines. Surface lipid metabolites quantified via GC-FID were averaged across biological replicates. Asterisks are listed above overexpression lines that are significantly different from the control according to individual t-tests performed for each metabolite ($p>0.05$). Means ± SE represent five (*cer4-3*) and three (T01) biological replicates from one transgenic event.

A.



B.

Figure 2 | Analysis of *zmGL1-like2* amino acid sequences (**A**) Phylogenetic tree of *zmGL1-like2* and homologous Arabidopsis and maize protein sequences based on maximum-likelihood. (**B**) Multiple sequence alignment of *zmGL1-like2* predicted protein products from cloned *Gl1-like2* transcripts and from GRMZM2G029912_T001 (B73 reference genome_V3) as well as Arabidopsis and maize homologs identified via BLAST. Boxed regions correspond to two putative conserved domains identified by NCBI DELTA-BLAST. Purple indicates the WAX2_C conserved domain and blue corresponds to the fatty acid hydroxylase conserved domain. Residues are displayed in bold and color-coded according to their chemical properties if they are < 9% conserved relative to the other aligned sequences (Gille, Fahling, Weyand, Wieland, & Gille, 2014).

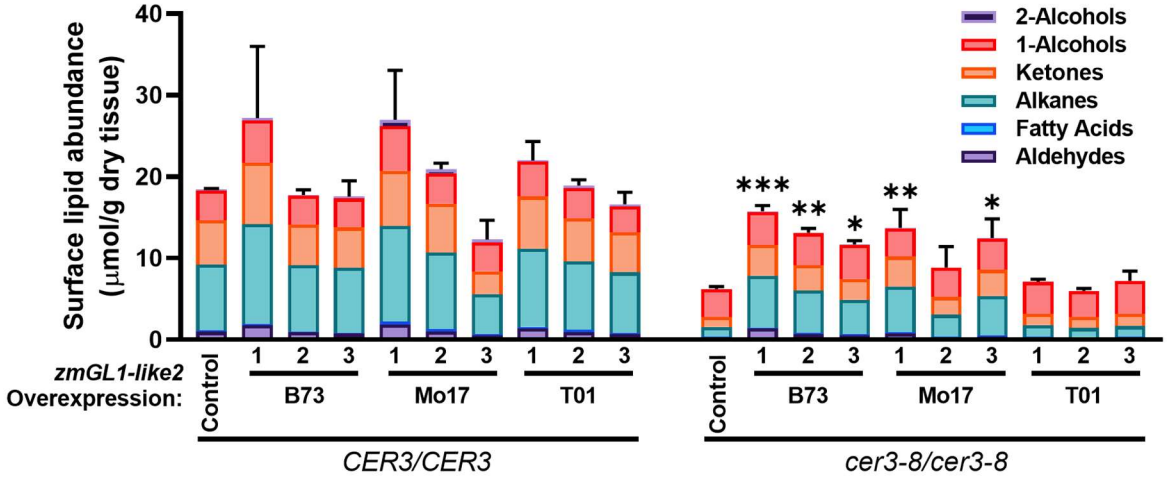


Figure 3 | Surface lipid composition on stems of Arabidopsis *cer3-8* and *CER3* lines overexpressing *zmGL1-like2* alleles. Three transformation events (1- 3) for each of the three *zmGl1-like2* alleles overexpressed within both the *CER3* wildtype (left) and *cer3-8* mutant (right) *Arabidopsis* lines. Means and SE are indicated for the total abundance of all lipid classes combined. Comparisons within the *cer3* and wildtype backgrounds were made between each transgenic event relative to the corresponding control using a Dunnett's test. Significant differences are indicated with asterisks ($p<0.05$*; $p<0.005$**; & $p<0.0005$***). Three biological replicates were assessed for each transgenic event with the exceptions of the control in the *CER3/CER3* background (n=2), the control in the *cer3-8/cer3-8* background (n=6), *ZmGl1-like2_Mo17* transgenic event #3 (n=2), and *ZmGl1-like2_T01* transgenic event #3 (n=2).

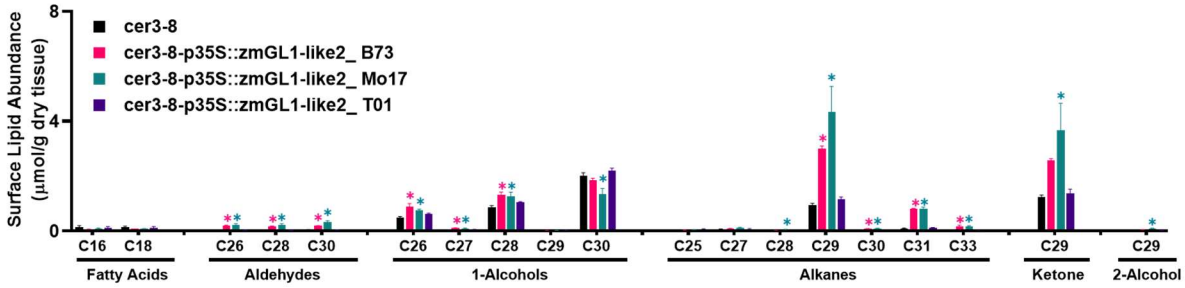


Figure 4 | Surface lipid composition on stems of Arabidopsis *cer3-8* mutant overexpressing *zmGl1-like2* alleles. Three *ZmGl1-like2* alleles were overexpressed within the *cer3-8* mutant Arabidopsis line. Asterisks are listed above overexpression lines that significantly differed from the control ($p>0.05$) according to individual Dunnett's tests performed for each metabolite. Means ± SE represent three biological replicates for each *ZmGl1-like2* allele and six biological replicates for the control. Results from one transgenic event for each allele are shown.

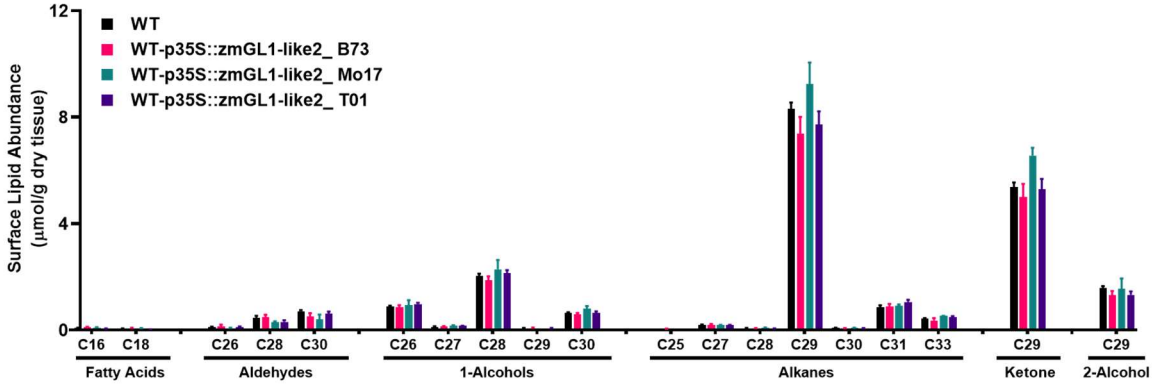
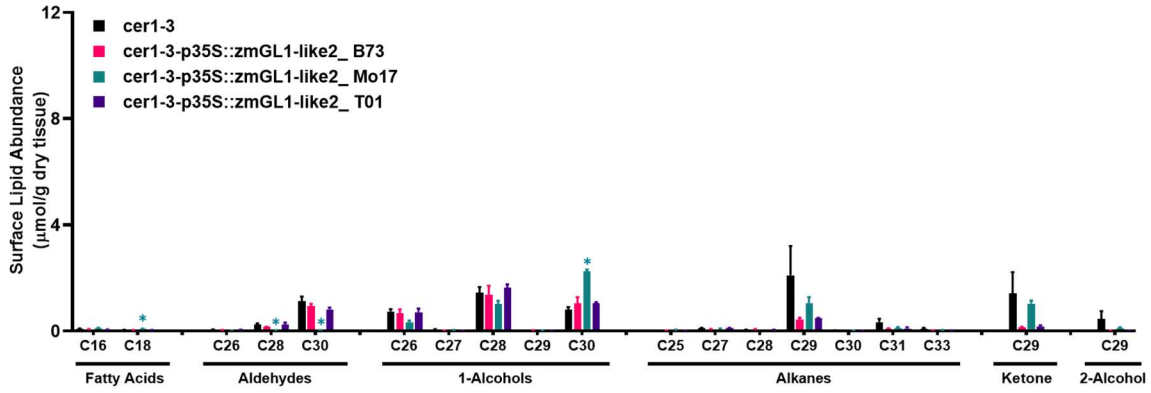


Figure 5 | Surface lipid composition on stems of wildtype Arabidopsis overexpressing *zmGl1-like2* alleles. Three *ZmGl1-like2* alleles were overexpressed within the wildtype Col-0 Arabidopsis line. Asterisks are not listed above overexpression lines because no significant differences were identified relative to the control ($p>0.05$) according to individual Dunnett's tests performed for each metabolite. Means ± SE; four biological replicates were assayed for *ZmGl1-like2_T01* and the wildtype control, five for *ZmGl1-like2_B73* and two for *ZmGl1-like2_Mo17*. Results from one transgenic event for each allele are shown.



[a]In two cases only two biological replicates were obtained.

Figure 6 | Surface lipid composition on stems of Arabidopsis *cer1-3* mutant overexpressing *zmGl1-like2* alleles. Three *ZmGl1-like2* alleles were overexpressed within the *cer1-3* mutant Arabidopsis line. Asterisks are listed above overexpression lines that significantly differed from the control ($p>0.05$) according to individual Dunnett's tests performed for each metabolite. Means ± SEM; three biological replicates were assayed for each *ZmGl1-like2* allele and eight biological replicates for the control. Results from one transgenic event for each allele are shown.

Supplement

Table S1. Amino acid similarity of GL1 and zmGL1-likes relative to Arabidopsis homologs.

			% Amino Acid Similarity to *Arabidopsis*		
Gene	**Organism**	**Gene Model**	**CER1**	**CER3**	**CER1-like**
ZmGlossy1	Maize	Zm00001d020557_T001	54	74	51
ZmGl1-Like1	Maize	Zm00001d015477_T001	53	74	49
ZmGl1-Like2	Maize	Zm00001d046865_T001	52	**76**	50
ZmGl1-Like3	Maize	Zm00001d051932_T002	70	31	66
ZmGl1-Like4	Maize	Zm00001d014055_T002	74	53	72
ZmGl1-Like5	Maize	Zm00001d017251_T003	72	53	72
CER1	Arabidopsis	AT1G02205.3	100	54	73
CER3	Arabidopsis	AT5G57800.1	54	100	53
CER1-like	Arabidopsis	AT1G02190.1	73	53	100

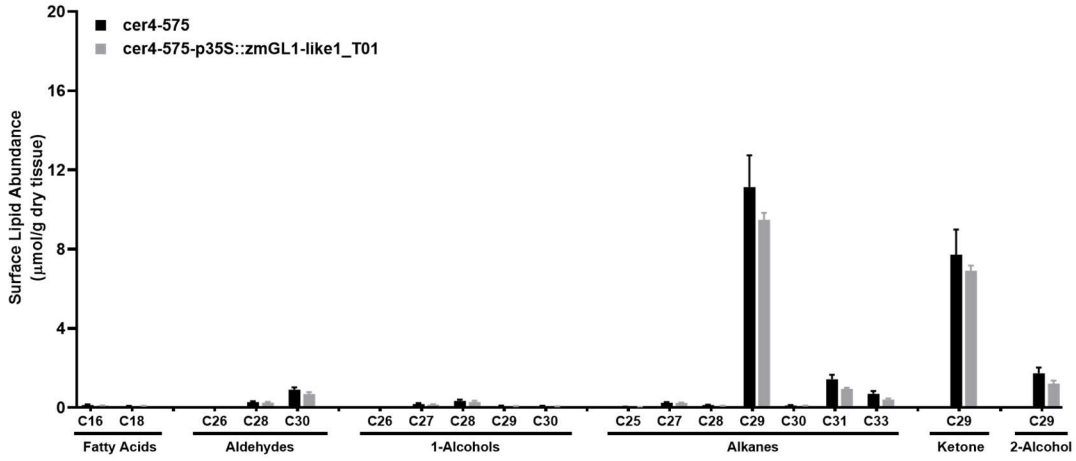


Figure S1 | Surface lipid composition on stems of Arabidopsis *cer4-575* mutant overexpressing *zmGl1-like1_T01*. *zmGl1-like2_T01* was overexpressed within the Arabidopsis *cer4-575* mutant line. No significant differences were found in overexpression lines relative to the control ($p>0.05$) according to individual Dunnett's tests performed for each metabolite. Means ± SE represent seven (control) and six (overexpression line) biological replicates from one transgenic event.

CHAPTER 5. CONCLUSIONS AND FUTURE DIRECTIONS

This dissertation sheds light on two facets of lipid metabolism within three biological systems, including *Saccharomyces cerevisiae*, *Zea mays*, and *Arabidopsis thaliana.* First, to increase our knowledge of genes that impact the accumulation of fatty acids and fatty acid derivatives in microbes, particularly triacylglycerols, putative causal mutations within *Saccharomyces cerevisiae* were identified through a forward genetic screen. Two resulting mutations were found to confer a lipid overaccumulation chemotype, thereby enhancing our knowledge of genes involved in lipid overproduction within yeast with potential applications in efforts to engineer oleochemicals. Second, to increase our knowledge of the roles of cuticular lipids in protecting plants against desiccation, and to better understand the functions of genes involved, associations between cuticular lipid composition and the rates at which water is lost from corn silks were evaluated, and a maize gene putatively involved in the biosynthesis of these cuticular metabolites, *ZmGl1-like2*, was functionally characterized in Arabidopsis.

In chapter two, an EMS mutagenesis screen was used to identify two point-mutations that contribute to fatty acid overaccumulation in *S. cerevisiae*. Of the seven mutations characterized from the lipid overaccumulation yeast screen, two caused increase lipid accumulation, primarily in the form of triacylglycerides. These mutations add to the knowledge-base that is currently being used to engineer microbial systems for oleochemical production (Pfleger, Gossing, & Nielsen, 2015; Y. J. Zhou et al., 2016). The *PIB2* (*phosphatidyl inositol binding2*) truncations discussed in Chapter 2 present a novel target for pathway engineering because they cause higher total fatty acid content (primarily in the form of triacylglycerols). Like the chemotypes found in the PIB2 truncations, the second causative point mutation identified in *TGL3* (*triacyl glycerol lipase3*) also caused elevated triacylglycerol accumulation. Interestingly, *TGL3* has already been

a successful target in rational design engineering approaches (Ferreira et al., 2018; Kurat et al., 2006) . The amino acid substitution identified in this study demonstrates that this amino acid is important for TGL3 function. In the future it would be enlightening to conduct biochemical studies of *tgl3-G289R* to determine why this amino acid substitution impairs function. For example, it could cause structural changes in protein folding or change the ability of TGL3 to interact with other proteins, metabolites, or membranes. Due to the residue location within the Patatin-like phospholipase domain (PFAM: PF01734) of TGL3, it seems most likely that this substitution impairs enzymatic function by directly disrupting the catalytic domain. Resolving the mode of causation for mutations that affect a given phenotype furthers our mechanistic understanding of these genes, which enables rational manipulation of these genes for downstream engineering applications.

Additional mutations were identified in the lipid over-accumulation forward genetic screen and can be explored in the future. For example, 151 sequence variants were identified within the population, 31 of which were nonsense or missense mutations consistent with EMS mutagenesis. Characterization of other mutations from this list of 151 sequence variants – including mutations within protein coding, regulatory, and/or intragenic regions – will likely reveal additional genes involved in lipid accumulation because the causal mutation(s) for some of the lipid overaccumulation strains identified in the EMS screen are still unidentified. Of the seven mutations that were assessed individually, two were tested in combination (i.e. mutations within both *TGL3* and *IRS4* that were identified in the *fao4* strain) because the single mutation in *TGL3* did not explain the entire lipid overaccumulation phenotype observed in the *fao4* strain. Other lipid overaccumulation strains identified within this screen may also have phenotypes that result from stacked mutations that could be further explored. Therefore, the list of 151 mutations

identified in this study are a resource that can be leveraged for identifying additional genes involved in lipid overaccumulation.

Chapter 3 assesses the agronomically-relevant trait of tolerance to water stress in plants (Hall, Chimenti, Trapani, Vilella, & de Hunau, 1984), particularly in relation to the plant cuticle and cuticular surface lipid composition. In this study, we identified a negative association between rates of water loss form maize silks with alkane abundance, a major lipid class within the silk extracellular surface metabolome (Dennison et al., 2019; Loneman et al., 2017; Perera et al., 2010). This association suggests a protective role of alkanes against desiccation stress. The role of alkenes in moderating water loss is also evidenced by previous work, including studies in saltwater cress and soybean (Xu et al., 2014; X. Zhou et al., 2014). A relationship between surface lipids and environmental factors, including relative humidity and temperature, were not identified in this study. However, natural variation in surface lipid compositions and their responsiveness to environmental factors such as water deficit and heat stress have been demonstrated (Dennison et al., 2019; Shepherd & Wynne Griffiths, 2006), suggesting that a larger panel of maize inbred lines with more variance across lipid traits may be necessary to discern these SL x E interactions. Alternatively, assessment of cuticular permeability (i.e. rates of water loss) in a mutant plant in which surface lipid biosynthesis and accumulation is impaired (e.g. the *cer1 Arabidopsis* mutant) would be the simplest means to test this relationship because it would limit variation due to other genetic factors that may also influence rates of water loss, such as morphology (Hannoufa et al., 1993).

Another major finding documented in Chapter 3 was the impact of husk-encasement status on rates of water loss from maize silks. Husk-encased silks were found to have lower rates of water loss than emerged silks, despite having equal to or lower alkane accumulation on the

silk surface. This suggests that other physical traits are more important in determining rates of water between the portions of silks that are emerged vs. encased by the husk leaves. Interestingly, emerged silks are known to have discontinuities in their cuticle, which could be reduced or nonexistent in husk-encased silks (Fuad-Hassan et al., 2008; Heslop-Harrison et al., 1985). Further investigation of discontinuities in emerged vs. husk-encased silks may explain these differences in cuticle permeability.

Finally, one might consider the impact of silk development on cuticular lipid accumulation and protection against water loss (Chapter 3). In an elegant study, Fuad-Hassan et al. (2008) described the overlapping gradients of cell division and cell expansion across the length of developing silks and determined that, near the point of emergence, silk growth relies on cell expansion as opposed to cell division. Therefore, initial silk water content relative to silk dry weight may reflect the degree of tissue expansion in emerged silks. Drought stress can delay silk emergence and reduce silk growth after silks emerge (Bassetti & Westgate, 1993; Fuad-Hassan et al., 2008; Herrero & Johnson, 1981). Therefore, emerged silk tissue with relatively low water content may correspond to cells which have not elongated after emergence. Cells which are recently elongated may have similar dry weights and surface lipid contents to unexpanded cells while having more exposed surface area to cover with a stretched out and thus thinner cuticular layer. Stunted growth under stressful conditions may serve to limit the susceptibility of plants to high rates of water loss. Future investigation into the cuticle thickness, surface cell size, rate of water loss, initial water content, and SL load from drought stressed silks relative to unstressed silk could aid in our understanding of relationships between silk elongation and potential changes in cuticle permeability.

In addition to the need to better understand how cuticle composition impacts cuticle function (Chapter 3), it is also important to uncover the specific functions of genes that govern the composition of these surface lipids. Chapter 4 demonstrated that *zmGL1-like2* is a functional homolog of *CER3*, a putative Acyl-CoA reductase in *Arabidopsis*. Moreover, a specific region of the primary sequence was shown to be critical for *zmGL1-like2* function. In a previous investigation, maize *Glossy1* and five *Glossy1*-like genes were expressed in Arabidopsis *cer1* and *cer3* mutants. All of the tested maize genes, with the exception of *ZmGl1-like2,* either fully or partially complemented either the Arabidopsis *cer1* or *cer3* mutants (Chmielowski, 2016). In contrast, the *ZmGl1-like2* construct did not complement either Arabidopsis mutant. In Chapter 4, we demonstrate that this lack of complementation was due to 13 additional amino acids (relative to all other homologs) that were inserted into the *zmGL1-like2_T01* sequence prior to the putative fatty acid hydroxylase domain of *CER3*. It is unclear why this 13-amino acid sequence prevents *zmGL1-like2_T01* to functionally complement *cer3* knockouts while the *ZmGl1-like2_B73* and *ZmGl1-like2_Mo17* alleles, which lack these 13 amino acids, were functional. This region may be important for protein association with CER1 or CER1-like, associations suggested by Bernard et al. and by Pascal et al., respectively (Bernard et al., 2012; Pascal et al., 2019). Another possibility is that this region affects tertiary structure either by interfering with protein folding or with its predicted interaction with the membrane (Berardini et al., 2015). Assessment of protein associations between *CER1* and *CER1-like* with *zmGL1-like2_T01* and *zmGL1-like2_B73* or *zmGL1-like2_Mo17* could shed light on whether the inserted 13 amino acids impair protein association.

Collectively, this work contributes to and enhances our understanding of lipid metabolism by identifying unique mutations that impact lipid outcomes, by corresponding

natural variation in plant cuticular surface lipid chemotypes to an agronomically important phenotypic outcome, and by confirming homology predicted gene function through complementation. Knowledge generated from this work has the potential to contribute to engineering efforts that target lipid metabolism for production of agricultural crops and bio-renewable products.

REFERENCES

Aarts, M. G., Keijzer, C. J., Stiekema, W. J., & Pereira, A. (1995). Molecular characterization of the CER1 gene of arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. *The Plant Cell*, *7*(12), 2115–2127. https://doi.org/10.1105/tpc.7.12.2115

Acree, W. E. J., & Chickos, J. S. (2018). Mass Spectra. In P. J. Linstrom & W. G. Mallard (Eds.), *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. https://doi.org/https://doi.org/10.18434/T4D303

Akaike, Y. (1985). Other oleochemical uses: Palm oil products. *Journal of the American Oil Chemists' Society*, *62*(2), 335–340. https://doi.org/10.1007/BF02541401

Akhtar, M. K., Turner, N. J., & Jones, P. R. (2013). Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Proceedings of the National Academy of Sciences*, *110*(1), 87–92. https://doi.org/10.1073/pnas.1216516110

Alberts B, Johnson A, L. J. (2002). *Molecular Biology of the Cell. 4th edition*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK26907/

Anderson, S. R., Lauer, M. J., Schoper, J. B., & Shibles, R. M. (2004). Pollination timing effects on kernel set and silk receptivity in four maize hybrids. *Crop Science*, *44*(2), 464–473. https://doi.org/10.2135/cropsci2004.0464

ASHRAE Research. (1993). *1993 ASHRAE Handbook: Fundamentals*.

Athenstaedt, K, & Daum, G. (2003). YMR313c/TGL3 encodes a novel triacylglycerol lipase located in lipid particles of Saccharomyces cerevisiae. *The Journal of Biological Chemistry*, *278*(26), 23317–23323.

Athenstaedt, K, Zweytick, D., Jandrositz, A., Kohlwein, S. D., & Daum, G. (1999). Identification and characterization of major lipid particle proteins of the yeast Saccharomyces cerevisiae. *Journal of Bacteriology*, *181*(20), 6441–6448.

Athenstaedt, Karin, & Daum, G. (2005). Tgl4p and Tgl5p, two triacylglycerol lipases of the yeast Saccharomyces cerevisiae are localized to lipid particles. *Journal of Biological Chemistry*, *280*(45), 37301–37309. https://doi.org/10.1074/jbc.M507261200

Avato, P., Bianchi, G., & Salamini, F. (1985). Absence of long chain aldehydes in the wax of the Glossy II mutant of maize. *Phytochemistry*, *24*(9), 1995–1997. https://doi.org/https://doi.org/10.1016/S0031-9422(00)83108-8

Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., … Zhao, H. (2015). Homology-Integrated CRISPR–Cas (HI-CRISPR) System for One-Step Multigene Disruption in Saccharomyces cerevisiae. *ACS Synthetic Biology*, *4*(5), 585–594. https://doi.org/10.1021/sb500255k

Barnes, J. D., Percy, K. E., Paul, N. D., Jones, P., McLaughlin, C. K., Mullineaux, P. M., … Wellburn, A. R. (1996). The influence of UV-B radiation on the physicochemical nature of tobacco (Nicotiana tabacum L.) leaf surfaces . *Journal of Experimental Botany*, *47*(1), 99–109. https://doi.org/10.1093/jxb/47.1.99

Bassetti, P., & Westgate, M. E. (1993). *Emergence, Elongation, and Senescence of Maize Silks*. *33*(19), 243–248.

Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., & Huala, E. (2015). The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. *Genesis*, *53*(8), 474–485. https://doi.org/10.1002/dvg.22877

Bernard, A., Domergue, F., Pascal, S., Jetter, R., Renne, C., Faure, J.-D., … Joubès, J. (2012). Reconstitution of plant alkane biosynthesis in yeast demonstrates that Arabidopsis ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *The Plant Cell*, *24*(7), 3106–3118. https://doi.org/10.1105/tpc.112.099796

Bianchi, A., Bianchi, G., Avato, P., & Salamini, F. (1985). *Biosynthetic pathways of epicuticular wax of maize as assessed by mutation, light, plant age and inhibitor studies*.

Bianchi, G., Avato, P., & Salamini, F. (1979). Glossy mutants of maize. *Heredity*, *42*(3), 391–395. https://doi.org/10.1038/hdy.1979.42

Bourdenx, B., Bernard, A., Domergue, F., Pascal, S., Léger, A., Roby, D., … Joubès, J. (2011). Overexpression of Arabidopsis ECERIFERUM1 promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiology*, *156*(1), 29–45. https://doi.org/10.1104/pp.111.172320

Brown, M., & Shanks, J. (2012). *Sustainable Bioenergy and Bioproducts*. 1–11. https://doi.org/10.1007/978-1-4471-2324-8

Cameron, K. D., Teece, M. A., & Smart, L. B. (2006). Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. *Plant Physiology*, *140*(1), 176–183. https://doi.org/10.1104/pp.105.069724

Campbell, A. A., Stenback, K. E., Flyckt, K., Hoang, T., Perera, M. A. D. N., & Nikolau, B. J. (2019). A single-cell platform for reconstituting and characterizing fatty acid elongase component enzymes. *PLOS ONE*, *14*(3), 1–18. https://doi.org/10.1371/journal.pone.0213620

Chan, R. K., & Otte, C. A. (1982). Physiological characterization of Saccharomyces cerevisiae mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. *Molecular and Cellular Biology*, *2*(1), 21–29. https://doi.org/10.1128/mcb.2.1.21

Chen, X., Goodwin, S. M., Boroff, V. L., Liu, X., & Jenks, M. A. (2003). Cloning and characterization of the WAX2 gene of Arabidopsis involved in cuticle membrane and wax production. *Plant Cell*, *15*(5), 1170–1185. https://doi.org/10.1105/tpc.010926

Chen, Y., Sheng, J., Jiang, T., Stevens, J., Feng, X., & Wei, N. (2016). Transcriptional profiling reveals molecular basis and novel genetic targets for improved resistance to multiple fermentation inhibitors in Saccharomyces cerevisiae. *Biotechnology for Biofuels*, *9*(1), 9. https://doi.org/10.1186/s13068-015-0418-5

Chmielowski, J. (2016). *The idenification of the GLOSSY1-Like gene family and their involvement in Zea mays surface lipid production*. Iowa State University.

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant Journal : For Cell and Molecular Biology*, *16*(6), 735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x

d'Espaux, L., Ghosh, A., Runguphan, W., Wehrs, M., Xu, F., Konzock, O., … Keasling, J. D. (2017). Engineering high-level production of fatty alcohols by Saccharomyces cerevisiae from lignocellulosic feedstocks. *Metabolic Engineering*, *42*, 115–125. https://doi.org/https://doi.org/10.1016/j.ymben.2017.06.004

de Jong, B. W., Shi, S., Valle-Rodriguez, J. O., Siewers, V., & Nielsen, J. (2015). Metabolic pathway engineering for fatty acid ethyl ester production in Saccharomyces cerevisiae using stable chromosomal integration. *Journal of Industrial Microbiology & Biotechnology*, *42*(3), 477–486. https://doi.org/10.1007/s10295-014-1540-2

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in Molecular Biology (Clifton, N.J.)*, *1151*, 165–188. https://doi.org/10.1007/978-1-4939-0554-6_12

Dennison, T., Qin, W., Loneman, D. M., Condon, S. G. F., Lauter, N., Nikolau, B. J., & Yandeau-Nelson, M. D. (2019). Genetic and environmental variation impact the cuticular hydrocarbon metabolome on the stigmatic surfaces of maize. *BMC Plant Biology*, *19*(1), 430. https://doi.org/10.1186/s12870-019-2040-3

Dickinson, W. J., Yang, Y., Schuske, K., & Akam, M. (1993). Conservation of Molecular Prepatterns during the Evolution of Cuticle Morphology in Drosophila Larvae. *Evolution*, *47*(5), 1396. https://doi.org/10.2307/2410155

Engel, S. R., Dietrich, F. S., Fisk, D. G., Binkley, G., Balakrishnan, R., Costanzo, M. C., … Cherry, J. M. (2014). The reference genome sequence of Saccharomyces cerevisiae: then and now. *G3 (Bethesda, Md.)*, *4*(3), 389–398. https://doi.org/10.1534/g3.113.008995

Ferreira, R., Teixeira, P. G., Gossing, M., David, F., Siewers, V., & Nielsen, J. (2018). Metabolic engineering of Saccharomyces cerevisiae for overproduction of triacylglycerols. *Metabolic Engineering Communications*, *6*, 22–27. https://doi.org/https://doi.org/10.1016/j.meteno.2018.01.002

Fonseca, A. E., & Westgate, M. E. (2005). Relationship between desiccation and viability of maize pollen. *Field Crops Research*, *94*(2–3), 114–125. https://doi.org/10.1016/j.fcr.2004.12.001

Fuad-Hassan, A., Tardieu, F., & Turc, O. (2008). Drought-induced changes in anthesis-silking interval are related to silk expansion: a spatio-temporal growth analysis in maize plants subjected to soil water deficit. *Plant, Cell & Environment*, *31*(9), 1349–1360. https://doi.org/10.1111/j.1365-3040.2008.01839.x

Gietz, R. D., & Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, *2*(1), 31–34. https://doi.org/10.1038/nprot.2007.13

Gille, C., Fahling, M., Weyand, B., Wieland, T., & Gille, A. (2014). Alignment-Annotator web server: rendering and annotating sequence alignments. *Nucleic Acids Research*, *42*(Web Server issue), W3-6. https://doi.org/10.1093/nar/gku400

Hall, A. J., Chimenti, C., Trapani, N., Vilella, F., & de Hunau, R. C. (1984). Yield in water-stressed maize genotypes: Association with traits measured in seedlings and in flowering plants. *Field Crops Research*, *9*, 41–57. https://doi.org/https://doi.org/10.1016/0378-4290(84)90005-4

Hannoufa, A., McNevin, J., & Lemieux, B. (1993). Epicuticular waxes of eceriferum mutants of Arabidopsis thaliana. *Phytochemistry*, *33*(4), 851–855. https://doi.org/https://doi.org/10.1016/0031-9422(93)85289-4

Herrero, M. P., & Johnson, R. R. (1981). Drought Stress and Its Effects on Maize Reproductive Systems1. *Crop Science*, *21*, 105–110. https://doi.org/10.2135/cropsci1981.0011183X002100010029x

Heslop-Harrison, Y., Heslop-Harrison, J., & Reger, B. J. (1985). THE POLLEN-STIGMA INTERACTION IN THE GRASSES. 7. POLLEN-TUBE GUIDANCE AND THE REGULATION OF TUBE NUMBER IN ZEA MAYS L. *Acta Botanica Neerlandica*, *34*(2), 193–211. https://doi.org/10.1111/j.1438-8677.1985.tb01879.x

Jarboe, L. R., Zhang, X., Wang, X., Moore, J. C., Shanmugam, K. T., & Ingram, L. O. (2010). Metabolic engineering for production of biorenewable fuels and chemicals: contributions of synthetic biology. *Journal of Biomedicine & Biotechnology*, *2010*, 761042. https://doi.org/10.1155/2010/761042

Javelle, M., Vernoud, V., Rogowsky, P. M., & Ingram, G. C. (2011). Epidermis: the formation and functions of a fundamental plant tissue. *The New Phytologist*, *189*(1), 17–39. https://doi.org/10.1111/j.1469-8137.2010.03514.x

Jenks, M. A., Rich, P. J., & Ashworth, E. N. (1994). Involvement of Cork Cells in the Secretion of Epicuticular Wax Filaments on Sorghum bicolor (L.) Moench. *International Journal of Plant Sciences*, *155*(5), 506. https://doi.org/10.1086/297190

Jenks, Matthew A., Tuttle, H. A., Eigenbrode, S. D., & Feldmann, K. A. (1995). Leaf epicuticular waxes of the Eceriferum mutants in Arabidopsis. *Plant Physiology*, *108*(1), 369–377. https://doi.org/10.1104/pp.108.1.369

Jetter, R., Kunst, L., & Samuels, A. L. (2006). Composition of plant cuticular waxes. *BIOLOGY OF THE PLANT CUTICLE*, *23*, 145–181. Retrieved from http://apps.webofknowledge.com.proxy.lib.iastate.edu/full_record.do?product=UA&search_mode=GeneralSearch&qid=21&SID=1DYWZtJu5Xr75FXgqcO&page=1&doc=4

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., … Ware, D. (2017). Improved maize reference genome with single-molecule technologies. *Nature*, *546*, 524. Retrieved from https://doi.org/10.1038/nature22971

Kamisaka, Y., Noda, N., Tomita, N., Kimura, K., Kodaki, T., & Hosaka, K. (2006). Identification of genes affecting lipid content using transposon mutagenesis in Saccharomyces cerevisiae. *Bioscience, Biotechnology and Biochemistry*, *70*(3), 646–653. https://doi.org/10.1271/bbb.70.646

Kass, R. E., & Raftery, A. E. (1995). Kass and Raftery 1995.pdf. *Jounral of the American Statistical Association*, Vol. 90, pp. 773–795.

Kates, M. (2010). *Techniques of Lipidology: Isolation, Analysis, and Identification of Lipids*.

Kerstiens, G. (1996). Cuticular water permeability and its physiological significance. *Journal of Experimental Botany*, *47*(12), 1813–1832. https://doi.org/10.1093/jxb/47.12.1813

Kim, A., & Cunningham, K. W. (2015). A LAPF/phafin1-like protein regulates TORC1 and lysosomal membrane permeabilization in response to endoplasmic reticulum membrane stress. *Molecular Biology of the Cell*, *26*(25), 4631–4645. https://doi.org/10.1091/mbc.E15-08-0581

Kim, K. S., Park, S. H., & Jenks, M. A. (2007). Changes in leaf cuticular waxes of sesame (Sesamum indicum L.) plants exposed to water deficit. *Journal of Plant Physiology*, *164*(9), 1134–1143. https://doi.org/10.1016/j.jplph.2006.07.004

Kim, K., Su, S., Park, H. D., Kim, K., & Jenks, M. A. (2007). Influence of Water Deficit on Leaf Cuticular Waxes of Soybean (Glycine max [L.] Merr.). *International Journal of Plant Sciences*, *168*(3), 307–316. https://doi.org/10.1086/510496

Kimbara, J., Yoshida, M., Ito, H., Kitagawa, M., Takada, W., Hayashi, K., … Ezura, H. (2013). Inhibition of CUTIN DEFICIENT 2 Causes Defects in Cuticle Function and Structure and Metabolite Changes in Tomato Fruit. *Plant & Cell Physiology*, *54*(9), 1535–1548. https://doi.org/10.1093/pcp/pct100

Kimura, K., Yamaoka, M., & Kamisaka, Y. (2004). Rapid estimation of lipids in oleaginous fungi and yeasts using Nile red fluorescence. *Journal of Microbiological Methods*, *56*(3), 331–338. https://doi.org/https://doi.org/10.1016/j.mimet.2003.10.018

Kolattukudy, P. E. (1985). Enzymatic Penetration of the Plant Cuticle by Fungal Pathogens. *Annual Review of Phytopathology*, *23*(1), 223–250. https://doi.org/10.1146/annurev.py.23.090185.001255

Kosma, D. K., Bourdenx, B., Bernard, A., Parsons, E. P., Lu, S., Joubes, J., & Jenks, M. A. (2009). The Impact of Water Deficiency on Leaf Cuticle Lipids of Arabidopsis. *Plant Physiology*, *151*(4), 1918–1929. https://doi.org/10.1104/pp.109.141911

Kosma, Dylan K, Bourdenx, B., Bernard, A., Parsons, E. P., Lü, S., Joubès, J., & Jenks, M. A. (2009). The impact of water deficiency on leaf cuticle lipids of Arabidopsis. *Plant Physiology*, *151*(4), 1918–1929. https://doi.org/10.1104/pp.109.141911

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Kurat, C. F., Natter, K., Petschnigg, J., Wolinski, H., Scheuringer, K., Scholz, H., … Kohlwein, S. D. (2006). Obese yeast: Triglyceride lipolysis is functionally conserved from mammals to yeast. *Journal of Biological Chemistry*, *281*(1), 491–500. https://doi.org/10.1074/jbc.M508414200

Kurata, T., Kawabata-Awai, C., Sakuradani, E., Shimizu, S., Okada, K., & Wada, T. (2003). The YORE-YORE gene regulates multiple aspects of epidermal cell differentiation in Arabidopsis. *Plant Journal*, *36*(1), 55–66. https://doi.org/10.1046/j.1365-313X.2003.01854.x

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25. https://doi.org/10.1186/gb-2009-10-3-r25

Le Provost, G., Domergue, F., Lalanne, C., Ramos Campos, P., Grosbois, A., Bert, D., … Gion, J.-M. (2013). Soil water stress affects both cuticular wax content and cuticle-related gene expression in young saplings of maritime pine (Pinus pinaster Ait). *BMC Plant Biology*, *13*(1), 95. https://doi.org/10.1186/1471-2229-13-95

Leber, C., & Da Silva, N. A. (2014). Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids. *Biotechnology and Bioengineering*, *111*(2), 347–358. https://doi.org/10.1002/bit.25021

Lee, M., Sharopova, N., Beavis, W. D., Grant, D., Katt, M., Blair, D., & Hallauer, A. (2002). Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. *Plant Molecular Biology*, *48*(5–6), 453–461. https://doi.org/10.1023/A:1014893521186

Lee, S. B., & Suh, M. C. (2015). Advances in the understanding of cuticular waxes in Arabidopsis thaliana and crop species. *Plant Cell Reports*, *34*(4), 557–572. https://doi.org/10.1007/s00299-015-1772-2

Leibundgut, M., Maier, T., Jenni, S., & Ban, N. (2008). The multienzyme architecture of eukaryotic fatty acid synthases. *Current Opinion in Structural Biology*, *18*(6), 714–725. https://doi.org/https://doi.org/10.1016/j.sbi.2008.09.008

Lennen, R. M., & Pfleger, B. F. (2013). Microbial production of fatty acid-derived fuels and chemicals. *Current Opinion in Biotechnology*, *24*(6), 1044–1053. https://doi.org/https://doi.org/10.1016/j.copbio.2013.02.028

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., … Ohlrogge, J. (2013). Acyl-lipid metabolism. *The Arabidopsis Book / American Society of Plant Biologists*, *11*, e0161. https://doi.org/10.1199/tab.0161

Li, L. I., Du, Y., He, C., Dietrich, C. R., Li, J., Ma, X., … Zheng, J. (2018). A novel maize gene, glossy6 involved in epicuticular wax deposition and drought tolerance. *BioRxiv*. https://doi.org/10.1101/378687

Liu, L., Pan, A., Spofford, C., Zhou, N., & Alper, H. S. (2015). An evolutionary metabolic engineering approach for enhancing lipogenesis in Yarrowia lipolytica. *Metabolic Engineering*, *29*, 36–45. https://doi.org/10.1016/j.ymben.2015.02.003

Liu, P., & Jarboe, L. R. (2012). Metabolic Engineering of Biocatalysts for Carboxylic Acids Production. *Computational and Structural Biotechnology Journal*, *3*(4), 1–9. https://doi.org/10.5936/csbj.201210011

Liu, T., Vora, H., & Khosla, C. (2010). Quantitative analysis and engineering of fatty acid biosynthesis in E. coli. *Metabolic Engineering*, *12*(4), 378–386. https://doi.org/https://doi.org/10.1016/j.ymben.2010.02.003

Lomakin, I. B., Xiong, Y., & Steitz, T. A. (2007). The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell*, *129*(2), 319–332. https://doi.org/10.1016/j.cell.2007.03.013

Loneman, D. M., Peddicord, L., Al-Rashid, A., Nikolau, B. J., Lauter, N., & Yandeau-Nelson, M. D. (2017). A robust and efficient method for the extraction of plant extracellular surface lipids as applied to the analysis of silks and seedling leaves of maize. *PLOS ONE*, *12*(7), 1–21. https://doi.org/10.1371/journal.pone.0180850

Long, L. M., Patel, H. P., Cory, W. C., & Stapleton, A. E. (2003). The maize epicuticular wax layer provides UV protection. *Functional Plant Biology*, *30*(1), 75. https://doi.org/10.1071/FP02159

Losev, E., Papanikou, E., OW, R., & BS, G. (2008). Cdc1p is an endoplasmic reticulum-localized putative lipid phosphatase that affects Golgi inheritance and actin polarization by activating Ca2+ signaling. *Molecular and Cellular Biology*, *28*(10), 3336–3343.

Lu, S., Zhao, H., Des Marais, D. L., Parsons, E. P., Wen, X., Xu, X., … Jenks, M. A. (2012). Arabidopsis ECERIFERUM9 Involvement in Cuticle Formation and Maintenance of Plant Water Status. *Plant Physiology*, *159*(3), 930–944. https://doi.org/10.1104/pp.112.198697

Madhavan, A., Jose, A. A., Binod, P., Sindhu, R., Sukumaran, R. K., Pandey, A., & Castro, G. E. (2017). Synthetic Biology and Metabolic Engineering Approaches and Its Impact on Non-Conventional Yeast and Biofuel Production. *Frontiers in Energy Research*, *5*(April). https://doi.org/10.3389/fenrg.2017.00008

Merchant, N., Lyons, E., Goff, S., Vaughn, M., Ware, D., Micklos, D., & Antin, P. (2016). The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. *PLOS Biology*, *14*(1), 1–9. https://doi.org/10.1371/journal.pbio.1002342

Moose, S. P., & Sisco, P. H. (1994). Glossy15 controls the epidermal juvenile-to-adult phase transition in maize. *Plant Cell*, *6*(10), 1343–1355. https://doi.org/10.1105/tpc.6.10.1343

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs*. Retrieved from http://cran.r-project.org/package=BayesFactor

Mortimer, R. K., & Johnston, J. R. (1986). Genealogy of principal strains of the yeast genetic stock center. *Genetics*, *113*(1), 35–43.

Moss, G. P., Smith, P. A. S., & Tavernier, D. (1995). Glossary of class names of organic compounds and reactive intermediates based on structure (IUPAC recommendations 1995). *Pure and Applied Chemistry*, Vol. 67, pp. 1307–1375. https://doi.org/10.1351/pac199567081307

Paidhungat, M., & Garrett, S. (1998). Cdc1 is required for growth and Mn2+ regulation in Saccharomyces cerevisiae. *Genetics*, *148*(4), 1777–1786.

Pascal, S., Bernard, A., Deslous, P., Gronnier, J., Fournier-Goss, A., Domergue, F., … Joubès, J. (2019). Arabidopsis CER1-LIKE1 Functions in a Cuticular Very-Long-Chain Alkane-Forming Complex. *Plant Physiology*, *179*(2), 415–432. https://doi.org/10.1104/pp.18.01075

Perera, M. A. D. N., Qin, W., Yandeau-Nelson, M., Fan, L., Dixon, P., & Nikolau, B. J. (2010). Biological origins of normal-chain hydrocarbons: a pathway model based on cuticular wax analyses of maize silks. *The Plant Journal : For Cell and Molecular Biology*, *64*(4), 618–632. https://doi.org/10.1111/j.1365-313X.2010.04355.x

Pfleger, B. F., Gossing, M., & Nielsen, J. (2015). Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabolic Engineering*, *29*, 1–11. https://doi.org/https://doi.org/10.1016/j.ymben.2015.01.009

Ploier, B., Scharwey, M., Koch, B., Schmidt, C., Schatte, J., Rechberger, G., … Daum, G. (2013). Screening for hydrolytic enzymes reveals Ayr1p as a novel triacylglycerol lipase in Saccharomyces cerevisiae. *The Journal of Biological Chemistry*, *288*(50), 36061–36072. https://doi.org/10.1074/jbc.M113.509927

Portwood, J. L. 2nd, Woodhouse, M. R., Cannon, E. K., Gardiner, J. M., Harper, L. C., Schaeffer, M. L., … Andorf, C. M. (2019). MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Research*, *47*(D1), D1146–D1154. https://doi.org/10.1093/nar/gky1046

Post-Beittenmiller, D. (1998). The cloned Eceriferum genes of Arabidopsis and the corresponding Glossy genes in maize. *Plant Physiology and Biochemistry*, *36*(1–2), 157–166. https://doi.org/10.1016/S0981-9428(98)80100-9

Premachandra, G. S., Saneoka, H., Kanaya, M., & Ogata, S. (1991). Cell membrane stability and leaf surface wax content as affected by increasing water deficits in maize. *Journal of Experimental Botany*, *42*(2), 167–171. https://doi.org/10.1093/jxb/42.2.167

R Core Team. (2019). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Retrieved from http://www.r-project.org/

Rajakumari, S., & Daum, G. (2010). Janus-faced enzymes yeast Tgl3p and Tgl5p catalyze lipase and acyltransferase reactions. *Molecular Biology of the Cell*, *21*(4), 501–510.

Riederer, M., & Schreiber, L. (2001). Protecting against water loss: analysis of the barrier properties of plant cuticles. *Journal of Experimental Botany*, *52*(363), 2023–2032. https://doi.org/10.1093/jexbot/52.363.2023

Rio, D. C., Ares, M. J., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5439. https://doi.org/10.1101/pdb.prot5439

Rowland, O., Lee, R., Franke, R., Schreiber, L., & Kunst, L. (2007). The CER3 wax biosynthetic gene from Arabidopsis thaliana is allelic to WAX2/YRE/FLP1. *FEBS Letters*, *581*(18), 3538–3544. https://doi.org/10.1016/j.febslet.2007.06.065

Rowland, O., Zheng, H., Hepworth, S. R., Lam, P., Jetter, R., & Kunst, L. (2006). CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. *Plant Physiology*, *142*(3), 866–877. https://doi.org/10.1104/pp.106.086785

Runguphan, W., & Keasling, J. D. (2014). Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. *Metabolic Engineering*, *21*, 103–113. https://doi.org/https://doi.org/10.1016/j.ymben.2013.07.003

Samuels, L., Kunst, L., & Jetter, R. (2008). Sealing plant surfaces: cuticular wax formation by epidermal cells. *Annual Review of Plant Biology*, *59*, 683–707. https://doi.org/10.1146/annurev.arplant.59.103006.093219

SAS Institute Inc. (2019). *JMP®, Version 14*. Cary, NC.

Schaeffer, M. L., Harper, L. C., Gardiner, J. M., Andorf, C. M., Campbell, D. a, Cannon, E. K. S., … Lawrence, C. J. (2011). MaizeGDB: curation and outreach go hand-in-hand. *Database : The Journal of Biological Databases and Curation*, *2011*, bar022. https://doi.org/10.1093/database/bar022

Schnable, P., Stinard, P., Wen, T., Heinen, S., Weber, D., Schneerman, M., … Nikolau, B. (1994). The genetics of cuticular wax biosynthesis. *Maydica*, *39*(4), 279–287. Retrieved from

http://apps.webofknowledge.com.proxy.lib.iastate.edu/full_record.do?product=UA&search_mode=GeneralSearch&qid=7&SID=1DYWZtJu5Xr75FXgqcO&page=2&doc=17

Schneider-Belhaddad, F., & Kolattukudy, P. (2000). Solubilization, partial purification, and characterization of a fatty aldehyde decarbonylase from a higher plant, Pisum sativum. *Archives of Biochemistry and Biophysics*, *377*(2), 341–349. https://doi.org/10.1006/abbi.2000.1798

Shepherd, T., & Wynne Griffiths, D. (2006). The effects of stress on plant cuticular waxes. *The New Phytologist*, *171*(3), 469–499. https://doi.org/10.1111/j.1469-8137.2006.01826.x

Shi, S., Ji, H., Siewers, V., & Nielsen, J. (2016). Improved production of fatty acids by Saccharomyces cerevisiae through screening a cDNA library from the oleaginous yeast Yarrowia lipolytica. *FEMS Yeast Research*, *16*(1), fov108. https://doi.org/10.1093/femsyr/fov108

Sitepu, I. R., Ignatia, L., Franz, A. K., Wong, D. M., Faulina, S. A., Tsui, M., … Boundy-Mills, K. (2012). An improved high-throughput Nile red fluorescence assay for estimating intracellular lipids in a variety of yeast species. *Journal of Microbiological Methods*, *91*(2), 321–328. https://doi.org/https://doi.org/10.1016/j.mimet.2012.09.001

Stein, S. E. (1999). An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *J Am Soc Mass Spectrom*, *10*(8), pp 770–781. https://doi.org/https://doi.org/10.1016/S1044-0305(99)00047-1

Sturaro, M., Hartings, H., Schmelzer, E., Velasco, R., Salamini, F., & Motto, M. (2005). Cloning and characterization of GLOSSY1, a maize gene involved in cuticle membrane and wax production. *Plant Physiology*, *138*(1), 478–489. https://doi.org/10.1104/pp.104.058164

Sukovich, D. J., Seffernick, J. L., Richman, J. E., Hunt, K. A., Gralnick, J. A., & Wackett, L. P. (2010). Structure, function, and insights into the biosynthesis of a head-to-head hydrocarbon in Shewanella oneidensis strain MR-1. *Applied and Environmental Microbiology*, *76*(12), 3842–3849. https://doi.org/10.1128/AEM.00433-10

Tanaka, H., Onouchi, H., Kondo, M., Hara-Nishimura, I., Nishimura, M., Machida, C., & Machida, Y. (2001). A subtilisin-like serine protease is required for epidermal surface formation in Arabidopsis embryos and juvenile plants. *DEVELOPMENT*, *128*(23), 4681–4689. Retrieved from http://apps.webofknowledge.com.proxy.lib.iastate.edu/full_record.do?product=UA&search_mode=GeneralSearch&qid=127&SID=1DYWZtJu5Xr75FXgqcO&page=1&doc=1

Thompson, C. J., Movva, N. R., Tizard, R., Crameri, R., Davies, J. E., Lauwereys, M., & Botterman, J. (1987). Characterization of the herbicide-resistance gene bar from Streptomyces hygroscopicus. *The EMBO Journal*, *6*(9), 2519–2523.

Treco, D. A., & Winston, F. (2008). Growth and manipulation of yeast. *Current Protocols in Molecular Biology*, *Chapter 13*, Unit 13.2. https://doi.org/10.1002/0471142727.mb1302s82

Varlakhanova, N. V, Mihalevic, M. J., Bernstein, K. A., & Ford, M. G. J. (2017). Pib2 and the EGO complex are both required for activation of TORC1. *Journal of Cell Science*, *130*(22), 3878–3890. https://doi.org/10.1242/jcs.207910

Venables, W.N., Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed.). New York, NY: Springer-Verlag.

Wakil, S. J., & Barnes, E. M. (1971). Fatty Acid Metabolism. In *Comprehensive Biochemistry*. https://doi.org/10.1016/B978-0-444-40950-8.50009-8

White, S. W., Zheng, J., Zhang, Y.-M., & Rock. (2005). The structural biology of type II fatty acid biosynthesis. *Annual Review of Biochemistry*, *74*, 791–831. https://doi.org/10.1146/annurev.biochem.74.082803.133524

Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., … Chen, Q.-J. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology*, *14*, 327. https://doi.org/10.1186/s12870-014-0327-y

Xu, X., Feng, J., Lü, S., Lohrey, G. T., An, H., Zhou, Y., & Jenks, M. a. (2014). Leaf cuticular lipids on the Shandong and Yukon ecotypes of saltwater cress, Eutrema salsugineum, and their response to water deficiency and impact on cuticle permeability. *Physiologia Plantarum*, *151*(4), 446–458. https://doi.org/10.1111/ppl.12127

Yang, S., Johnston, N., Talideh, E., Mitchell, S., Jeffree, C., Goodrich, J., & Ingram, G. (2008). The endosperm-specific ZHOUPI gene of Arabidopsis thaliana regulates endosperm breakdown and embryonic epidermal development. *Development (Cambridge, England)*, *135*(21), 3501–3509. https://doi.org/10.1242/dev.026708

Yeats, T. H., & Rose, J. K. C. (2013). The formation and function of plant cuticles. *Plant Physiology*, *163*(1), 5–20. https://doi.org/10.1104/pp.113.222737

Zhou, X., Jenks, M. A., Liu, J., Liu, A., Zhang, X., Xiang, J., … Chen, X. (2014). Overexpression of Transcription Factor OsWR2 Regulates Wax and Cutin Biosynthesis in Rice and Enhances its Tolerance to Water Deficit. *Plant Molecular Biology Reporter*, *32*(3), 719–731. https://doi.org/10.1007/s11105-013-0687-8

Zhou, X., Li, L., Xiang, J., Gao, G., Xu, F., Liu, A., … Wan, X. (2015). OsGL1-3 is involved in cuticular wax biosynthesis and tolerance to water deficit in rice. *PLoS ONE*, *10*(1). https://doi.org/10.1371/journal.pone.0116676

Zhou, Y. J., Buijs, N. A., Zhu, Z., Qin, J., Siewers, V., & Nielsen, J. (2016). Production of fatty acid-derived oleochemicals and biofuels by synthetic yeast cell factories. *Nature Communications*, *7*(1), 11709. https://doi.org/10.1038/ncomms11709